

STIC-Biotech/ChemLib

119478

From: Hutzell, Paula
Sent: Thursday, April 15, 2004 11:11 AM
To: Graser, Jennifer; STIC-Biotech/ChemLib
Subject: RE: rush search

RECEIVED
APR 15 2004
STIC-BIOTECH/CHMLIB
(STIC)

approved

-----Original Message-----

From: Graser, Jennifer
Sent: Thursday, April 15, 2004 10:03 AM
To: Hutzell, Paula
Subject: rush search
Importance: High

Hi Paula,

Could you please authorize the following rush search for an amendment which is due?

Thanks,
Jennifer

STIC:

Please search SEQ ID NOS: 1; 1 (101-1176); 2; 3; 4; 5; 6; 6 (101-853); 6 (853-1176) from **10/608,504** in pending and commercial databases. They are species of one another and I need all of them searched.

Thanks,
Jennifer Graser
REMSEN
EO3 B09/ 3C18 mailbox
AU 1645
272-0858

Searcher: D. Schweitzer
Phone: 272-2526
Location: Remsen EO1 A61
Date Picked Up: _____
Date Completed: 4/20
Searcher Prep/Review: 15
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 7
AA Sequences: 2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: Compuser
WWW/Internet: _____
Other (specify): _____



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



Email: Raf.Podowk@ogr.ki.se
Class: NotI site

FEATURES

Location/Qualifiers
1. 473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human NotI clones"

ORIGIN

Query Match 4.3%; Score 54.6; DB 28; Length 473;
Best Local Similarity 56.2%; Pred. No. 0.0036;
Matches 99; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

293 TGGGAGGCGCCAGCTTTTCCGGCTGATTTCCGGGCTCCACGAAATGCTGATC 352
283 TGGAGGCGCGAGGCTCTCGCCATCGTCGCGGCGCGCGCAGTGTGCGCATC 342
353 GCCCTGTTGGGCGCAGCGCCCTGGCGCCATCGCGCTCACCAATGCTGTGAA 412
343 GGCATGCTTAAGGCGGCGCCCACTGCTGATCGTGAACCACTGCTGCTGACT 402
413 TTCCGCGCAGTATCTGATGCTTTTATTCCTCGCTGATGCTGATAAAACCCCT 468
403 TCGCAGACCTGCTTACGCTTGTCCATGCGCGCTGCTGATACGAGCACCCTT 458

RESULT 2
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION BX356664 Clone CS0D1015B03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D1015CA02NP1.

FEATURES

source

1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015B03"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.5%; Score 45; DB 13; Length 1201;
Best Local Similarity 12.9%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 47; Conservative 166; Mismatches 148; Indels 0; Gaps 0;

307 GTTTCGCGGCTGATTTTCGCGGCTCCACGAAATGCTGATGCGCCCTGTTGTGG 366
745 ST 804

367 CGAGCGCCCTGCGGCGCCATCGGCTCACCATGCTGATGTAATTCGCCAGATT 426
805 SBTCTSSSSSSSBTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 864
427 CTATGCGTTTATTCGCGCTGATGCTCAAAAACCCATGTCCTTATTCGCT 486
865 YSBSST 924
487 TTTCGCGCTTATGAGAACCTAGCGAGTCACTGCGCGAGCGCGAGCTGTGCGC 546
925 BTSSSBST 984
547 GTGCGCATTAATCTCAATGCAATAGCGTTTCACTTCACTGATGCTGCGGCTTC 606
985 VTSSSBST 1044
607 CGAGTGGCATGCGCAGAGTGAATTCCTTTGAAATTAAGGCGCTGAGTGGCCTTG 666
1045 SSSBSST 1104
667 CTC 669
1105 TTS 1107

RESULT 3
BX376097 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX376097 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION BX376097 cDNA clone CS0C022YM12 5-PRIME, mRNA sequence.
ACCESSION BX376097
VERSION BX376097.1 GI:30434756
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2866.f
Contact: Feng Liang Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D022B06QPI.

FEATURES

source

1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D022YM12"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.3%; Score 41.8; DB 13; Length 1201;
Best Local Similarity 9.1%; Pred. No. 2.7;
Matches 55; Conservative 238; Mismatches 310; Indels 2; Gaps 1;

593 TTGCGGCTCTACCGAGTGGCATGCGAGATTATCTTTGAAATTAAGGCGCTC 652
441 TTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 500
653 GAGTTCGCGCTTCTCTCTCTTGTGACGCTGACTTGTGATTCCTGCGAAGCAAG 712

[illegible]

RESULT 4	BX381961	1201 bp	mRNA	linear	EST 08-MAY-2003
LOCUS	BX381961				
DEFINITION	BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI072YF05 3-PRIME, mRNA sequence.				
ACCESSION	BX381961				
VERSION	BX381961.1	GI:30453007			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Pojaves, D.				
TITLE	1. (bases 1 to 1201)				
JOURNAL	Full-length cDNA libraries and normalization				
COMMENT	Unpublished (2001)				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	BP 191 91006 Evry cedex - France				
	Email: seqest@genoscope.cns.fr, Web : www.genoscope.cns.fr				
	Library was constructed by Life Technologies, a division of				
	Invitrogen. Contact : Peng Liang Email : liliang@lifetech.com URL :				
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600				
	Paradise Avenue Genoscope sequence ID : CSODI072CC03NPL.				
FEATURES					
source	1..1201				
	Location/Qualifiers				

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GSD01072XP05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/tissue_type="Homo sapiens PLACENTA COT 25-NORMALIZED"
/clone="1st strand cDNA was primed with a NotI-oligo(dT

```

ORIGIN

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

	Query Match	Similarity	3.2%; Score 40.8; DB 13; Length 1201;	
	Best Local	Similarity	3.0%; Pred. No.5/2;	
	Matches	Conservative	191; Mismatches 362; Indels 0; Gaps 0;	
Oy	397	CACATTGCTGTGAACATCCGCCAGCATATCTAAGCGTTTCATCCCGCGCAGATGGCT	456	
Db	406	CKNNNNNNKKKKNCKNNKKKKMMKKMMKKKKKKKKKKKKNNNNNCNANNNNNKKKKNNNN	465	
Oy	457	CAAAAACCATTGCCCGGTTTCTATTGGTTTTGCAGCTTATCGAGAAGCTTAGCGAGT	516	
Db	466	NNNTKNNNNNNCNKKNNKKYKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	525	
Oy	517	CACGCGGCCAGCGCCGCGAGGCTGTGGGGGTGCGACTTATCTCAATGCAATAGCGTT	576	
Db	526	MNNNNCNMMAANNKKNNKKNNKKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	585	
Oy	577	TCACTCCTACTGTGGTAATTCGGCGGTCTCACCGAGTGCGAGATCGAGAATTATCTCTT	636	
Db	586	NN	645	
Oy	637	TGAATTAAGGGCTCGAGTTCGGCTTGGCTCTCTTTGTACAAGCTTAGATTTCAGATTC	696	
Db	646	KNN	705	
Oy	697	CTGCCGAAAGAAAAAGAGATCCCTTCTCTGCTGCTCGAGTTGAGCTTACACATTCG	756	
Db	706	NKKKKKTNNKKKKNNNAKNNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	765	
Oy	757	TCTTGGTGTAATTCAGATGAGCGGCTATTTCGCGCGCTGTGATCTTCTGTGGTCTGT	816	
Db	766	KNN	825	
Oy	817	GACCATCCGGTACTTCTTTTGGAAAAGGCTGCTAAATGACAATGACTGATTTCTCCTGAT	876	
Db	826	KCKCMCBMMCKVCMMCKMCBKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	885	
Oy	877	CTCCTTGTTCGCGATATGACGATGACTATTGCTTCGCTCGCGGGCGGTTCCGTTCTTA	936	
Db	886	MMKKNNMMBKKNCMKWKJNCMKWKKKKKKKKNNNMKKMMMKCKBMCMCKNNNNNNNNNN	945	
Oy	937	ATCCTTAAGCCCCCTAGCGTAATCAATTT	966	
Db	946	MMNBKCVMKKMKMKCKKVMNBKCKACKKK	975	

RESULT 5
BX540413/c 692 bp mRNA linear EST 28-JUL-2003
LOCUS
DEFINITION
ACCESSION BX540413
VERSION BX540413
KEYWORDS
SOURCE
ORGANISM
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 692).
Herault, F., De Meuth-Netzinger, V., Desert, C., Recout, E., Plumi, F.,
Klopp, C. and Douaire, M.
Construction and primary characterization of chicken normalized
multi-tissue cDNA libraries
unpublished (2003)
Contact: Douaire M
INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63

Fax: +33 (0) 2.23.48.54.70

Email: Madeleine.Douaire@roazhon.inra.fr
 Sequence cleaned of vector, adaptor and repetitions. Contact us
 at sigensaupport@jouy.inra.fr to obtain the chromatogram of this
 sequence.
 Plate: 0001 row: d column: 10
 Seq primer: M13R

FEATURES

source

Location/Qualifiers

1..692
 /organism="Gallus gallus"
 /mol_type="rRNA"
 /db_xref="taxon:9031"
 /clone="gscab0001b.d.10"
 /issue_type="adipose tissue, granulosa, multi-tissues,
 oviduct, small follicle, utero-vaginal gland"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="AGENAE Gallus gallus multi-tissues library
 (gscab)"
 /note="Vector: pTTT3D-pac; Clone distribution: AGENAE
 Resource centre, Francois Piumi, Francois Piumi.inra.fr,
 INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine
 de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 3.1%; Score 39.4; DB 13; Length 692;

Best Local Similarity 62.9%; Pred. No. 10; Mismatches 36; Indels 0; Gaps 0;

Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTGGGCTCTGTGACCATCCGCTACTTCTTTGGGAAAGCTGCTAATGCACTGAT 864
 DB 321 CTTGGGCGAGTTGGCCATCTGCTCTGCACTTTSCACAAACGACGAGACGCGCAT 262
 QY 865 TTCTCCTGATTCCTCTGTTGTCGCAATGTCGAG 901
 DB 261 TTTGCTAAGCTCACCCTTCTTCTGTCGACCTCTGAG 225

RESULT 6

AJ397466/c

LOCUS AJ397466 718 bp mRNA linear EST 25-JAN-2001
 DEFINITION AJ397466 dkfz426 Gallus gallus cDNA clone 217r1, mRNA sequence.
 ACCESSION AJ397466
 VERSION AJ397466.1 GI:7130433
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 718)
 Abdurakhmanov, I., Lodygin, D., Gerolt, P., Arakawa, H., Law, A.,
 Plachy, J., Korn, B. and Buerstedde, J.M.
 A large database of chickenursal ESTs as a resource for the
 analysis of vertebrate gene function
 Genome Res. 10 (12), 2062-2069 (2000)
 20568495
 11116100

COMMENT

Contact: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinistr. 52, 20251 Hamburg, Germany
 Email: URL: http://genome.hpi.uni-hamburg.de/dt0est.html.
 Location/Qualifiers
 1..718
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="CB"
 /db_xref="taxon:9031"
 /clone="217r1"
 /issue_type="Bursa of Fabricius"
 /cell_type="Bursal lymphocyte"
 /dev_stage="2-3 weeks old"

FEATURES

source

ORIGIN

/clone_lib="dkfz426"

Query Match 3.1%; Score 39.4; DB 9; Length 718;

Best Local Similarity 62.9%; Pred. No. 10; Mismatches 36; Indels 0; Gaps 0;

Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTTGGCTCTGTGACCATCCGCTACTTCTTTGGGAAAGCTGCTAATGCACTGAT 864
 DB 117 CTTGGGCGAGTTGGCCATCTGCTCTGCACTTTSCACAAACGACGAGACGCGCAT 58
 QY 865 TTCTCCTGATTCCTCTGTTGTCGCAATGTCGAG 901
 DB 57 TTTGCTAAGCTCACCCTTCTTCTGTCGACCTCTGAG 21

RESULT 7

CC684922

LOCUS CC684922 743 bp DNA linear GSS 19-JUN-2003
 DEFINITION OGUAX55TH ZM 0.7-1.5 KB zea mays genomic clone ZMMBMA0397J14,
 genomic survey sequence.
 ACCESSION CC684922
 VERSION CC684922.1 GI:32089698
 KEYWORDS GSS.
 SOURCE Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 743)
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGUAX55TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES

source

Location/Qualifiers
 1..743
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0397J14"
 /clone_lib="ZM_0.7-1.5 KB"
 /note="Vector: pBSK-; Site 1: HincII, 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 3.1%; Score 39.4; DB 29; Length 743;

Best Local Similarity 53.6%; Pred. No. 11; Mismatches 71; Indels 0; Gaps 0;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 531 CCGCAGGCTGTGTCGGGTCGACCTATATCAATGCAATAGGCTTCACTCTACTGGG 590
 DB 475 CTTGCGCTCCGCGCGGTCGTCCTCCACATCTCTATCTCGGATGGCAATTCATAG 534
 QY 591 TATTGCGGCTCTCAACCGAGTGGCGATGSCAAGTTGATTCCTTTGAATTAAGGCGC 650
 DB 535 GTTTCGGCGGCGGCGGATCCAGTCCGCTTCGCGAGCGCTTAAGCTTCAGGTC 594
 QY 651 TCGAGTTCGCCCTTTGCTCTCTCTTTGTACGCG 683
 DB 595 ACGAGAGGCGGCTTTGCTCTCTCCCTGTCGCCG 627

RESULT 8
BU448643/c
LOCUS
DEFINITION 60321136F1 CSFORBN13 Gallus gallus cDNA clone CHEST191c7 5', mRNA
756 bp mRNA linear EST 29-NOV-2002

ACCESSION BU448643
VERSION BU448643
KEYWORDS GI:25937954
SOURCE EST
ORGANISM Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 756)
AUTHORS Boardman, P.E., Sans-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNA
Curt. Biol. 12 (22), 1965-1969 (2002)
2235534

TITLE JOURNAL
MEDLINE
PUBMED 12445392

COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 756
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST191c7"
/sex="Female"
/dev_stage="adult"
/lab_host="MDH10B"
/clone_11b="CSFORBN13"
/note="Organ: ovary; Vector: Bluescript II KS(+); Site: 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntend, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
Bluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1996) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reaming hybridization was used."

ORIGIN
Query Match 3.1%; Score 39.4; DB 13; Length 756;
Best Local Similarity 62.9%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTTGGGCTCTGACACCCGCTACTCTCTTGGGAAAGGCTGTAATGACACAT 864
DB 270 CTTGGGCGAGTGGGCATCTGCTCTGACCTTTCACACGACGACGTGACCCGAT 211
QY 865 TTCTCCTGATCTCTCTGTTGTCGACATATGTGAG 901
DB 210 TTTTGTAAAGTCACCTTCTGCTGAGCCTCTGAG 174

RESULT 9
CNS009WA/c
LOCUS
DEFINITION CNS009WA 787 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR20K04 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.
ACCESSION AL054013
VERSION AL054013.1 GI:4935590
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygora;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 787)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Ooegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. 787
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR20K04"
/clone_11b="RPCI-98"
/note="end : 17"

ORIGIN
Query Match 3.1%; Score 39.4; DB 29; Length 787;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 76; Conservative 15; Mismatches 76; Indels 0; Gaps 0;

QY 652 CGAGTGGCCCTTGTCTCTCTTGTCAACGCTTGAATTCCTGCCAGCAAAA 711
DB 334 CTACTTGCTCTCTCTCTCTTGTGTATAMKMKWCTTGAATCCCTKTVAGARAAAC 275
QY 712 GCAGATCCCTCTCTCTCTCTCTGACAGTTGACCTTCAACATTCCTGTGTAATCC 771
DB 274 AYTGTGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 215
QY 772 AGGTCAAGCCCTATTTGCGGCGCTGATCTCTTGGGCTGTGA 818
DB 214 TAATTCGCATATGTTATATGATTTTATTTATGTCGCGCTGA 168

RESULT 10
BX425603
LOCUS
DEFINITION BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
C10B0222A07 3-PRIME, mRNA sequence.
885 bp mRNA linear EST 15-MAY-2003

ACCESSION BX425603
VERSION BX425603.1 GI:30770486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 885)
Full-length cDNA libraries and normalization
Unpublished (2001)

Page 8

		3.0%;	Score 38;	DB 14;	Length 835;
		Best Local Similarity 51.1%;	Pred. No. 28;		
Matches	89;	Conservative	0;	Mismatches	85;
				Indels	0;
				Gaps	0;
QY	352	CGCGCTCGTTGTGGGCGCAGCGCCCTCGGGGCGCCATCGCGCTACACACATTGCTGTAA	411		
Db	79	CGACCGCGCGCTCCGGGTGCGGGCGTGTGGCGCTTTCGCGCGGTGGCGAGATGGCTTC	138		
QY	412	CTTCGCGCACGATCTCTATGCGTTTATTCCGCTGCATGTGATCAAAAACCCCATTCG	471		
Db	139	ATGCGCGCGCGGTCTTCACCGGAATGGCGACCCGACCTGCGGGCTGGAGAGCGGTGCT	198		
QY	472	CCGATTCTATTTCGTTTTGCGCGCTTATCGACAGAAAGCTTACGCACTACCTCGCGG	525		
Db	199	CTCCGCTCTACGCGCGCTCTTCGTGGCAGAGAGAGTGGCCTGGCCACCTCTCGCGG	252		

Search completed: April 17, 2004, 20:01:40

Job time : 4123.78 secs

Search completed: April 17, 2004, 20:01:40
Job time : 4123.78 secs

```

Query Match: 3.0%; Score 38; DB 14; Length 835;
Best Local Similarity: 51.1%; Pred. No. 28;
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 352 CGCCTCGTTGTGGGCGCAGCGCCCTGGGGCCCATGCGCTACACACATTGCTGTGAA 411
Db 79 CGACGCGGCTCTCGGTGCGCGGCTGCGCGTCTTGCGCGGCTGCGGAGATGCTTC 138
QY 412 CTTCGCGCAGCATTCATGCGTTCCTCCGCTGCATGTGATCAAAAACCCCATTTGC 471
Db 139 ATCCGCGCGGCTTCGACCGGATCGCGCACCCGACCTGCGGCTTGGAAGCGCTGCT 198
QY 472 CCGTTTCATTCGATTTCGCGCTTATCGACGAAAGCTTACGCACTGCGCGC 525
Db 199 CTCGCGCTACGCGCGTCTTCGTGCGCAGGGAAGTGGCTGCGCACCTTCGCGGC 252

Search completed: April 17, 2004, 20:01:40
Job time : 4123.78 secs

```

```
/note=Vector: pBluescript II KS +; Site_1: EcoRI; Site_2
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
```


Tue Apr 20 06:47:12 2004

us-10-608-504-1.rge

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 5387 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271

Sequence: 1 ggcgcataacagatcctag.....agcctaacgacagcgcct 1271

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_da.*
2: gb_hlg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hlg_hum.*
31: em_hlg_inv.*
32: em_hlg_other.*
33: em_hlg_mus.*
34: em_hlg_pin.*
35: em_hlg_rtd.*
36: em_hlg_mam.*
37: em_hlg_vtl.*
38: em_sy.*
39: em_hlgo_hum.*
40: em_hlgo_mus.*
41: em_hlgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	6 AR391953	AR391953 Sequence
2	1271	100.0	1271	6 AR391956	AR391956 Sequence
3	1271	100.0	1271	6 AX137709	AX137709 Sequence
4	1271	100.0	1271	6 AX137714	AX137714 Sequence
5	1271	100.0	1271	6 BD014991	BD014991 Nucleotide
6	1271	100.0	1271	6 BD014994	BD014994 Nucleotide
7	1271	100.0	340000	6 AP005274	AP005274 Coryneb
8	1271	100.0	349980	6 AX120085	AX120085 Sequence
9	1260	99.1	2105	1 AF454053	AF454053 Coryneb
10	753	59.2	753	6 AR391954	AR391954 Sequence
11	753	59.2	753	6 AX120373	AX120373 Sequence
12	753	59.2	753	6 AX137710	AX137710 Sequence
13	753	59.2	753	6 BD014992	BD014992 Nucleotide
14	753	59.2	753	6 BD162490	BD162490 Nucleotide
15	753	59.2	397	6 AX137083	AX137083 Sequence
16	365.6	28.8	397	6 AX137085	AX137085 Sequence
17	365.6	28.8	324	6 AR391955	AR391955 Sequence
18	324	25.5	324	6 AX120374	AX120374 Sequence
19	324	25.5	324	6 AX137712	AX137712 Sequence
20	324	25.5	324	6 BD014993	BD014993 Nucleotide
21	324	25.5	324	6 BD162491	BD162491 Nucleotide
22	185	14.6	715	6 AX137077	AX137077 Sequence
23	185	14.6	715	6 BD013995	BD013995 Novel nuc
24	149.6	11.8	87340	1 AP005224	AP005224 Coryneb
25	132.8	10.4	53793	1 BX248361	BX248361 Coryneb
26	127.4	10.0	711	6 AX123492	AX123492 Sequence
27	127.4	10.0	711	6 BD165609	BD165609 Novel pol
28	127.4	10.0	309400	6 AX127153	AX127153 Sequence
29	127.4	10.0	325651	1 AP005283	AP005283 Coryneb
30	59.8	4.7	349926	1 BX571660	BX571660 Nucleotide
31	58.6	4.6	756	6 AR386287	AR386287 Sequence
32	57.2	4.5	9888	1 AE011050	AE011050 Methanosa
33	54.4	4.3	856	9 HSA325448	AJ325448 Homo sapi
34	52.4	4.1	7218	6 H64994	166494 Sequence
35	51.8	4.1	11025	1 AE013585	AE013585 Methanosa
36	51.2	4.0	125020	9 AF429315	AF429315 Homo sapi
37	51	4.0	308015	1 AE016783	AE016783 pseudomon
38	48.2	3.8	2000	6 AX655393	AX655393 Sequence
39	48	3.8	10732	6 E32986	E32986 Gene encodi
40	46.6	3.8	301838	1 AE017209	AE017209 Geobacter
41	46.6	3.7	3619	1 ECH410307	AJ410307 Erwinia c
42	46.6	3.7	10733	1 AE013695	AE013695 Yersinia
43	46.6	3.7	220050	1 AJ414156	AJ414156 Yersinia
44	46	3.6	215050	1 AL646084	AL646084 Ralstonia
45	45.8	3.6	20342	1 AE008888	AE008888 Salmonella

ALIGNMENTS

RESULT 1
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kemmerknecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;

FEATURES

```
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"
```

ORIGIN

Query Match	100.0%;	Score 1271;	DB 6;	Length 1271;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1271; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	GGCGAGTCATGGAAATCTAGCTTCATATATATGGCAAAATAGCTGATGGAGTGGCGAAAC	60
Qy	1	GGCGAGTCATGGAAATCTAGCTTCATATATATGGCAAAATAGCTGATGGAGTGGCGAAAC	60
Db	1	GGCGAGTCATGGAAATCTAGCTTCATATATATGGCAAAATAGCTGATGGAGTGGCGAAAC	60
Qy	61	TGGCAACAAAACCTACCCGGGCAATTGTGTGATGATTTGTATAGTGGCAAAAACGCAAGAT	120
Db	61	TGGCAACAAAACCTACCCGGGCAATTGTGTGATGATTTGTATAGTGGCAAAAACGCAAGAT	120
Qy	121	TCATTCAGAGCTGAGAGTGTCCCATCCAGGACGCCCTGGAAACGATGATTAAGGTTA	180
Db	121	TCATTCAGAGCTGAGAGTGTCCCATCCAGGACGCCCTGGAAACGATGATTAAGGTTA	180
Qy	121	TCATTCAGAGCTGAGAGTGTCCCATCCAGGACGCCCTGGAAACGATGATTAAGGTTA	180
Db	121	TCATTCAGAGCTGAGAGTGTCCCATCCAGGACGCCCTGGAAACGATGATTAAGGTTA	180
Qy	181	TGGGCGCTACGAAATGGCGCAAGTCTAAAACTCCCTTGCTGCAGGTTTGGGCAATGTA	240
Db	181	TGGGCGCTACGAAATGGCGCAAGTCTAAAACTCCCTTGCTGCAGGTTTGGGCAATGTA	240
Qy	241	CCGATATGGTATATGCGTATTGTCTCTGGTATATCAATACGAGCTACGAATGTGTGGGCAAC	300
Db	241	CCGATATGGTATATGCGTATTGTCTCTGGTATATCAATACGAGCTACGAATGTGTGGGCAAC	300
Qy	301	CCCACTGTTTTCCGGGCTGATTTTTGGCGGGCTCCACGGAATGCTGTGATCCGCTCGT	360
Db	301	CCCACTGTTTTCCGGGCTGATTTTTGGCGGGCTCCACGGAATGCTGTGATCCGCTCGT	360
Qy	361	TGTGGGCGACAGGCCCTCGGCGGCATCGCGCTCACCAATTTGCTGGTGAATCTTCGCCCA	420
Db	361	TGTGGGCGACAGGCCCTCGGCGGCATCGCGCTCACCAATTTGCTGGTGAATCTTCGCCCA	420
Qy	421	CGTATTCATGCGTTTTTCATTTCCGCGTGCATGTGTCAAAAACCCATTCGCCGTTTCTA	480
Db	421	CGTATTCATGCGTTTTTCATTTCCGCGTGCATGTGTCAAAAACCCATTCGCCGTTTCTA	480
Qy	481	TTGGGTTTTGCGCCTTATGACAGGAAGCCTACGACATCTCGCGCCACAGGCCGCAAGCTG	540
Db	481	TTGGGTTTTGCGCCTTATGACAGGAAGCCTACGACATCTCGCGCCACAGGCCGCAAGCTG	540
Qy	541	GTCGCGCTGCGCATTTATCTCAATGCAATAACGTTTACTCTCACTAGGATATCGGCGG	600
Db	541	GTCGCGCTGCGCATTTATCTCAATGCAATAACGTTTACTCTCACTAGGATATCGGCGG	600
Qy	601	TCTCACCGAGATGGCGATGCGAAGTTGATTTCTCTTTGAATTTAAGGCGCTCGAGTTGCG	660
Db	601	TCTCACCGAGATGGCGATGCGAAGTTGATTTCTCTTTGAATTTAAGGCGCGCTCGAGTTGCG	660
Qy	661	CCTTTGCTCTCTTTGTACGCTGACTTTGATTTCCGCGAAGCAAAAAGAGATCCC	720
Db	661	CCTTTGCTCTCTTTGTACGCTGACTTTGATTTCCGCGAAGCAAAAAGAGATCCC	720
Qy	721	TTCTCTGCTGCTCGACAGTTTGAAGTTCAACATCTCTTGTGTAATTCAGGCTCAGGC	780
Db	721	TTCTCTGCTGCTCGACAGTTTGAAGTTCAACATCTCTTGTGTAATTCAGGCTCAGGC	780
Qy	781	CCATATTGGCGGCGTGTGATCTCTTTGGGTCGTGTGACATCCGGTACTCTCTTGAG	840
Db	781	CCATATTGGCGGCGTGTGATCTCTTTGGGTCGTGTGACATCCGGTACTCTCTTGAG	840
Qy	841	AAAGGCTGCTAAATGACACTGATTTCTCCTGATTTCTCTTGTGTGACAGTATGTGA	900
Db	841	AAAGGCTGCTAAATGACACTGATTTCTCCTGATTTCTCTTGTGTGACAGTATGTGA	900
Qy	901	GTCATTAATTTTGGGCTCCGGGCGGTTCCGTTCTTAATCTTAAACCCCTACGATATCA	960
Db	901	GTCATTAATTTTGGGCTCCGGGCGGTTCCGTTCTTAAATCTTAAACCCCTACGATATCA	960

QY	961	CAATTGTGGGCAAAAATGGCGATGTGGATCCGACAGAAATCCCTGCAATTGGACGCA	1020
Db	961	CAATTGTGGGCAAAAATGGCGATGTGGATGCGACAGAAATCCCTGCAATTGGACGCA	1020
QY	1021	TCAACGTTTGGCAGCAATGCGATNGATCTGAAGACTCTAACCTTTGGTCTCATTTGCCGT	1080
Db	1021	TCAACGTTTGGCAGCAAAATGCGATNGATCTGAAGACTCTAACCTTTGGTCTCATTTGCCGT	1080
QY	1081	GCGATTACAGTGTGGGCGCATCTTCTTGGCGGTGACGCACTTGTGAGCGTTGGCGCT	1140
Db	1081	GCGATTACAGTGTGGGCGCATCTTCTTGGCGGTGACGCACTTGTGAGCGTTGGCGCT	1140
QY	1141	GGCACCATCGCTTTGTGTGACCTGGTGAAATCTTTCTTAAACCTGCAATTAATTA CAATAAT	1200
Db	1141	GGCACCATCGTTTTTGTGTGACCTGGTGAAATCTTTCTTAAACCTGCAATTAATTA CAATAAT	1200
QY	1201	CCGCATGCGCTCAATTGGAAGGGGATGCGGATTTTTTAAAGAACCTTGAAGAAAGGCTTAAG	1260
Db	1201	CCGCATGCGCTCAATTGGAAGGGGATGCGGATTTTTTAAAGAACCTTGAAGAAAGGCTTAAG	1260
QY	1261	CAGACAGCGCT 1271	
Db	1261	CAGACAGCGCT 1271	

RESULT 2

LOCUS	AR391356	1271 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 6 from patent US 6613345.				
ACCESSION	AR391356				
VERSION	AR391356.1	GI:40115727			

ORIGIN

Query Match	100.0%;	Score 1271;	DB 6;	Length 1271;
Best Local Similarity	100.0%;	Pred. Nc. 0;		
Matches 1271; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	GCGGCACTGAAGGAACTC	TAGCTCTAATAATATG	CAACAAATGCTCTAGTTAGAGTGCGCAAC	60
Db	1	GCGCATCAATGGAACTC	TAGCTCTAATAATATG	CAACAAATGCTCTAGTTAGAGTGCGCAAC	60
QY	61	TGGCAACAAAAC	TACCCGGCAATTGTGTG	ATGATTTTACTGTGCAAAAACGCAAGAAAT	120
Db	61	TGGCAACAAAAC	TACCCGGCAATTGTGTG	ATGATTTTACTGTGCAAAAACGCAAGAAAT	120
QY	121	TCATTTCAAGCCTG	GAGGTGTGCGCATCCAAAGC	CAGCCTCGAACCAGATGATTAAGGTTA	180
Db	121	TCATTTCAAGCCTG	GAGGTGTGCGCATCCAAAGC	CAGCCTCGAACCAGATGATTAAGGTTA	180
QY	181	TGGCGGCTACGAAAT	TGGCGCAAGGCTCTAAAACCTC	CTCTGTGACAGTTTGGCGCATGA	240
Db	181	TGGCGGCTACGAAAT	TGGCGCAAGGCTCTAAAACCTC	CTCTGTGACAGTTTGGCGCATGA	240
QY	241	CCCGATTGGTATTGG	CTTGTGCTCTCTGGTATTCAATAC	GGCTCTCGAATGGTGGCGAGC	300
Db	241	CCCGATTGGTATTGG	CTTGTGCTCTCTGGTATTCAATAC	GGCTCTCGAATGGTGGCGAGC	300
QY	301	CCCACTGTTTCCG	GCCTGATTTTTCGCGG	CTCCACGGAATGCTGGTCAATGCGCCTGCT	360
Db	301	CCCACTGTTTCCG	GCCTGATTTTTCGCGG	CTCCACGGAATGCTGGTCAATGCGCCTGCT	360

Db 301 CCCACTGTTTCCGCGCTGATTTTCGCGGCGCTCCACCGAATGCTGTCCTCCCTCCT 360
 Qy 361 TGTGGGGGCGAGCGCGCCCTGGGCGCGCATCGCGCTCGACCACTGCTGGTGAATTCGCGCA 420
 Db 361 TGTGGGGGCGAGCGCGCCCTGGGCGCGCATCGCGCTCGACCACTGCTGGTGAATTCGCGCA 420
 Qy 421 CGATTCATGCGCTTTTCATTCCTCGCTGATGTTGTGCAAAAACCCCATGCGCTTTCTA 480
 Db 421 CGATTCATGCGCTTTTCATTCCTCGCTGATGTTGTGCAAAAACCCCATGCGCTTTCTA 480
 Qy 481 TTTGGTTTGGCGCTTATTCAGACGCTACGCTACGCTGCGCGCGCGCGCGCGCGCTG 540
 Db 481 TTTGGTTTGGCGCTTATTCAGACGCTACGCTACGCTGCGCGCGCGCGCGCGCGCTG 540
 Qy 541 GTGCGCGCGCGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 600
 Db 541 GTGCGCGCGCGCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 600
 Qy 601 TCTCAACGAGAGTGGCGATCGCAGAGTTGATTCCTTTGAAATTAAGGCGCTCGATTCG 660
 Db 601 TCTCAACGAGAGTGGCGATCGCAGAGTTGATTCCTTTGAAATTAAGGCGCTCGATTCG 660
 Qy 661 CCTTGGCTCTCTTTGTCAGCTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTT 720
 Db 661 CCTTGGCTCTCTTTGTCAGCTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTT 720
 Qy 721 TTTCTGCTGCTGCGAGGTTGAGCTTCACCTGATTCCTTTGATTCCTTTGATTCCTTT 780
 Db 721 TTTCTGCTGCTGCGAGGTTGAGCTTCACCTGATTCCTTTGATTCCTTTGATTCCTTT 780
 Qy 781 CCTATTGGCGCGCTGCTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTTGATTC 840
 Db 781 CCTATTGGCGCGCTGCTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTTGATTC 840
 Qy 841 AAAGGCTGCTAAATGACAACTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTT 900
 Db 841 AAAGGCTGCTAAATGACAACTGATTCCTTTGATTCCTTTGATTCCTTTGATTCCTTT 900
 Qy 901 GTCAATTACTTTGGCGCTCGGCGCGCTTCCTTTGATTCCTTTGATTCCTTTGATTCCT 960
 Db 901 GTCAATTACTTTGGCGCTCGGCGCGCTTCCTTTGATTCCTTTGATTCCTTTGATTCCT 960
 Qy 961 CAATTGCGCGCAAAATGGCGATGTCAGAGGAGATTCCTTTGATTCCTTTGATTCCTTT 1020
 Db 961 CAATTGCGCGCAAAATGGCGATGTCAGAGGAGATTCCTTTGATTCCTTTGATTCCTTT 1020
 Qy 1021 TCAACGTTTGGCGAGATGCGATGATTCGAAAGCTTCAACCTTTGATTCCTTTGATTC 1080
 Db 1021 TCAACGTTTGGCGAGATGCGATGATTCGAAAGCTTCAACCTTTGATTCCTTTGATTC 1080
 Qy 1081 GCGATTACGATGAGTGGCGCATCTTTGCGCGTTCAGCGACCTTTGATTCCTTTGATTC 1140
 Db 1081 GCGATTACGATGAGTGGCGCATCTTTGCGCGTTCAGCGACCTTTGATTCCTTTGATTC 1140
 Qy 1141 GCGACCATGCTTTTGTGGAAGTGTGATCTTTTCTAAAACCTGATTAATAAATAAAT 1200
 Db 1141 GCGACCATGCTTTTGTGGAAGTGTGATCTTTTCTAAAACCTGATTAATAAATAAAT 1200
 Qy 1201 CCGCATGCGCTCAATTTGAAGGAGTGGGATTTTGAAGACCTTAATAAAGGCTTAAG 1260
 Db 1201 CCGCATGCGCTCAATTTGAAGGAGTGGGATTTTGAAGACCTTAATAAAGGCTTAAG 1260
 Qy 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

RESULT 3
 LOCUS AX137709 1271 bp DNA linear PAT 30-MAY-2001
 ACCESSION Sequence 1 from Patent EP1096010.
 AX137709
 VERSION AX137709.1 GI:14273886

KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.
 REFERENCE
 AUTHORS Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 TITLE Nucleotide sequences coding for branched-chain amino acids export
 JOURNAL Proteins, method for isolating them and their use
 Patent: EP 1096010-A 1 02-MAY-2001;
 Degussa AG (DE) / FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 FEATURES
 source
 1. /1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="EMBL:CA39923"
 /translation="MKTQEHSSLEVSPEKSKALBPDDKGYRVEIAGLKTSLAAGL
 GNPYIGIARGILVIGYEWMAPLPSGLIFAGSTEMVIALVGAAPLGAIALTL
 VNFNRYFAFSPLIVVKNRPFYSVPAIDEAAYTRAPRAGSARWLLSMOIAFH
 SYWFGQLGVAIAELIPEIKGELPSFVLTLIDSCRTKQIPLILLAGLSFTI
 ALVVI PQGLFPAALIFLGLTIRYFEGKAK"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39924.1"
 /db_xref="GI:14273888"
 /db_xref="EMBL:CA39924"
 /translation="MTDFSCILLVAVCAVTFALRAVPFLILPEBSQFVGKAM
 WNPAGIALITLTFSPRSNAIDIKTLTFLIILAVALIVVAHLLOGRRTLLSVGATIVF
 GLVNLFF"
 ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCGCGATCAATGGAATCTAGCTTCATATATTTGACAAATGCGCTAGTGTGCGCAAC 60
 Db 1 GCGCGATCAATGGAATCTAGCTTCATATATTTGACAAATGCGCTAGTGTGCGCAAC 60
 Qy 61 TGGCAACAAATCTACCGGCAATGTGTGATGATTTAGTGTGCAAAAACCGAAGAT 120
 Db 61 TGGCAACAAATCTACCGGCAATGTGTGATGATTTAGTGTGCAAAAACCGAAGAT 120
 Qy 121 TCAATTAAGCGTGGAGTGTGCGATTCGCAAGGAGCGCTGGAACGAGATTAAGGTTA 180
 Db 121 TCAATTAAGCGTGGAGTGTGCGATTCGCAAGGAGCGCTGGAACGAGATTAAGGTTA 180
 Qy 181 TCGGCGCTAGGAATGCGCAAGTCTAAAAACCTTCCTGCTGCAAGTTTGGCATGTA 240
 Db 181 TCGGCGCTAGGAATGCGCAAGTCTAAAAACCTTCCTGCTGCAAGTTTGGCATGTA 240
 Qy 241 CCGGATTTGATTTGCTTTGCTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 300
 Db 241 CCGGATTTGATTTGCTTTGCTCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTTGCTTT 300
 Qy 301 CCGACTGTTTCCGCGCTGATTTTCGCGGAGCTCCACCGAATGCTGTCATGCGCTCGT 360
 Db 301 CCGACTGTTTCCGCGCTGATTTTCGCGGAGCTCCACCGAATGCTGTCATGCGCTCGT 360
 Qy 361 TGTGGGGGCGAGCGCGCCCTGGGCGCGCATCGCGCTGACCACTGCTGTGATTCGCGCA 420
 Db 361 TGTGGGGGCGAGCGCGCCCTGGGCGCGCATCGCGCTGACCACTGCTGTGATTCGCGCA 420

QY	421	CGAATTCATAGCGGTTTCATTCGCCGCTGCAGATGAGGTCAAAAACCCCATTCGCCGTTTCA	480
Db	421	CGAATTCATAGCGGTTTCATTCGCCGCTGCAGATGAGGTCAAAAACCCCATTCGCCGTTTCA	480
QY	481	TTCCGATTTTCGCGCTTAATCGACGAAAGCTTAACGAGTCATCTCGGCCACAGCCCGCAGGCTTG	540
Db	481	TTCCGATTTTCGCGCTTAATCGACGAAAGCTTAACGAGTCATCTCGGCCACAGCCCGCAGGCTTG	540
QY	541	GTCCGCGCTGGGAGATTATCTCAATGCAAAATAGCGTTTCACTCCCTCACTGGGATTTCCGCGCG	600
Db	541	GTCCGCGCTGGGAGATTATCTCAATGCAAAATAGCGTTTCACTCCCTCACTGGGATTTCCGCGCG	600
QY	601	TCTCAGCGGAGATGGCGCATTCGACAGAGTTGATTCCTTTGAAATTAAGGCGCTCGAGTTTCG	660
Db	601	TCTCAGCGGAGATGGCGCATTCGACAGAGTTGATTCCTTTGAAATTAAGGCGCTCGAGTTTCG	660
QY	661	CCTTTGCTCTCTCTTTTGTCAGCGTCGACTTTTGAGATTTCTCTCGCGAAACGAAAAGCAGATCCC	720
Db	661	CCTTTGCTCTCTCTTTTGTCAGCGTCGACTTTTGAGATTTCTCTCGCGAAACGAAAAGCAGATCCC	720
QY	721	TTCTCTGCTGCTCGAGAGTTTGAGGTTTCAACCATGCTCTTGTGTAATTCAGGTCAGGC	780
Db	721	TTCTCTGCTGCTCGAGAGTTTGAGGTTTCAACCATGCTCTTGTGTAATTCAGGTCAGGC	780
QY	781	CCTAATTTGCGGCGCTGCTGATCTTCTTGAGGTCTGTGACCATCCGGTAATTCTTCTTTGGG	840
Db	781	CCTAATTTGCGGCGCTGCTGATCTTCTTGAGGTCTGTGACCATCCGGTAATTCTTCTTTGGG	840
QY	841	AAAGGCTGCTAAATGACAACTGATTTCCTCGTATTCCTCTGTGTGTCGAGATGATGCA	900
Db	841	AAAGGCTGCTAAATGACAACTGATTTCCTCGTATTCCTCTGTGTGTCGAGATGATGCA	900
QY	901	GTCAATTACTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTGAATCA	960
Db	901	GTCAATTACTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTGAATCA	960
QY	961	CAATTTTGGGAGGCAAAATGGCCATGTGCATATGCGACAGAGAAATCCTTGGCCATTTTGACCGCA	1020
Db	961	CAATTTTGGGAGGCAAAATGGCCATGTGCATATGCGACAGAGAAATCCTTGGCCATTTTGACCGCA	1020
QY	1021	TCAAAGTTTCGACGAATGCGATAGATGGAAGACTTAACTTAACCTTTGTGTCATTTGCCGTT	1080
Db	1021	TCAAAGTTTCGACGAATGCGATAGATGGAAGACTTAACTTAACCTTTGTGTCATTTGCCGTT	1080
QY	1081	GCGATTTACAGTGTGTGGCGCATTTCTTGCGGCTCAACGCACTTGTGTAGAGGTTGGCGCT	1140
Db	1081	GCGATTTACAGTGTGTGGCGCATTTCTTGCGGCTCAACGCACTTGTGTAGAGGTTGGCGCT	1140
QY	1141	GGCAGCATATGTTTTTGTGTGGAATGAGATCTTTCTTAAACTGAGTAAATACAAAAT	1200
Db	1141	GGCAGCATATGTTTTTGTGTGGAATGAGATCTTTCTTAAACTGAGTAAATACAAAAT	1200
QY	1201	CCGCAATGCCCTCAATTTGAAAGGGAGATGCGAATTTTTTAAAGAACTTAAAGAAAGGCTTAAG	1260
Db	1201	CCGCAATGCCCTCAATTTGAAAGGGAGATGCGAATTTTTTAAAGAACTTAAAGAAAGGCTTAAG	1260
QY	1261	CAGACAGCGCT 1271	
Db	1261	CAGACAGCGCT 1271	

RESULT 4	AX137714	1271 bp	DNA	linear	PAT 30-MAY-2001
LOCUS	AX137714				
DEFINITION	Sequence 6 from Patent EP106010.				
ACCESSION	AX137714				
VERSION	AX137714.1	GI:14273893			
KEYWORDS					
SOURCE					
ORGANISM	<i>Corynebacterium glutamicum</i>				
	<i>Corynebacterium glutamicum</i>				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacteriaceae; Corynebacterium.				
REFERENCE	1				

AUTHORS Kernenknecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
TITLE Nucleotide sequences coding for branched-chain amino acids export proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 6 02-May-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES Location/Qualifiers
1
1021

SOURCE

CDS

CDS

ORIGIN

Query Match	100.0%	Score 1271;	DB 6;	Length 1271;
Best Local Similarity	100.0%	Prod. No. 0;		
Matches 1271;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0

1 GCGGATCATGGATCTAGCTTCATATTTGCACATAGCCAGTGGAGTGGCAAC 80

D5 1 GCGCGATCAATGGATCTAGCTTCATATATGGCCTATAGCCATGATGAGGAGCCCTTTC

61 TGGCAACAAATAC TACCCTGTA TGTGTATGTAT TGATCATCCTT CTTCCCTTCCT

[illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

240

101 TCCCTCCTTGGGCAATGTA 240

241 CCGGATTGCTATTGCGTTTGGTCTCTTGGTTATTCAATACGGCTACGAATGGTGGCAGC 300

Db 241 CCCGATGGTATTCGCGTTGGTCTCTTGGTATTCATACGGCTACGAATGGTGGCAGC 300

301 CCCACTGTTTCCGGCCTGATTTTCGGGGCTCCACCGAATGCTGTCAATGCCCTCGT 360

Db 301 CCGACTGTTTCCGGGCTGATTTCGGGGCTCCACCGAATGCTGGTCATCGCCCTCGT 360

361 TGTGGCGCAGCGCCCTGGGCGGCATCGCGCTCACCAATTGCTGTGAATTCCGCCA 420

Db 361 TGTGGCGCAGCCCCCTGGCGGCATCGCGCTCACCCACATTGCTGTGTACTTCCGCCA 420

421 CGATTCTATGCGTTTTCATTCCCGCTGCATGTGGTCAAAACCCCATGGCCGTTCTA 480

D6 421 CGTATTCCTAAGCGTTTCAATCCCGCAGAGGGACAAACCACAATGCCCCGATTC

481

1260
1260
2002
1d
W. J. [Signature]

Query Match	100.0%	Score 1271;	DB 6;	Length 1271;
Best Local Similarity	100.0%	Prod. No. 0;		
Matches 1271;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	1	GCGGCATCAATGGAATCTAGCTTCATATATTTGCAATATGCTGATAGGTGCGCAAC	60	
Qy	61	TGGCAACAAACCTACCCGGCAATGTGTGAGATGTAAGTGTGCAAAAAGCGAAGAT	120	
Db	61	TGGCAACAAACCTACCCGGCAATGTGTGAGATGTAAGTGTGCAAAAAGCGAAGAT	120	
Qy	121	TCATTCAAGCTGTGAGGTGTGCGCATCCAAAGCGACCTGTGAAACGATGATTAAGTTA	180	
Db	121	TCATTCAAGCTGTGAGGTGTGCGCATCCAAAGCGACCTGTGAAACGATGATTAAGTTA	180	
Qy	181	TCGGCGCTACGAATTCGGCGCAAGCTCTAAAACTCCCTTGTGTGAGGTTTGGGATGTA	240	
Db	181	TCGGCGCTACGAATTCGGCGCAAGCTCTAAAACTCCCTTGTGTGAGGTTTGGGATGTA	240	
Qy	241	CCGATGTTGATTTGGGTTTGGTCTCTGTTGATTAATACGAGCTACGAATGGTGGCGAGC	300	
Db	241	CCGATGTTGATTTGGGTTTGGTCTCTGTTGATTAATACGAGCTACGAATGGTGGCGAGC	300	
Qy	301	CCCACTGTTTCCGGCTGATTTTTCGGCGGCTCCAACGAATGCTGTGATTCGCCCTGT	360	
Db	301	CCCACTGTTTCCGGCTGATTTTTCGGCGGCTCCAACGAATGCTGTGATTCGCCCTGT	360	
Qy	361	TGTGGGCGACAGCCCGCTGGGGCGCATGCGCTACACATGTCGTGTAATTCGCCCA	420	
Db	361	TGTGGGCGACAGCCCGCTGGGGCGCATGCGCTACACATGTCGTGTAATTCGCCCA	420	
Qy	421	CGATATTCATGCGTTTTCATTCCCGCTGCATGTGTCAAAAAACCCATTGCCGTTTCTA	480	
Db	421	CGATATTCATGCGTTTTCATTCCCGCTGCATGTGTCAAAAAACCCATTGCCGTTTCTA	480	
Qy	481	TTTGGTTTTCGCGCTTATCGAAGAGCTTACGAGTCACTGCGGCGCAAGCCGCGAGGCT	540	
Db	481	TTTGGTTTTCGCGCTTATCGAAGAGCTTACGAGTCACTGCGGCGCAAGCCGCGAGGCT	540	
Qy	541	GTCGGCGTGGGAGCTTATCTCAATGCAAAATAGCGTTTCACTCTCTAGTATTCGGCGG	600	
Db	541	GTCGGCGTGGGAGCTTATCTCAATGCAAAATAGCGTTTCACTCTCTAGTATTCGGCGG	600	
Qy	601	TTTCAACCGGAGTGGCGATTCGCAAGATTGATCTTTTGAATTAAGGGCTTGGAGTTGCC	660	
Db	601	TTTCAACCGGAGTGGCGATTCGCAAGATTGATCTTTTGAATTAAGGGCTTGGAGTTGCC	660	
Qy	661	CGTTGGCTCTCGCTTTGTCAAGCTAGCTTGGATTCCTGCGCAACGAAAAACAGATGCC	720	
Db	661	CGTTGGCTCTCGCTTTGTCAAGCTAGCTTGGATTCCTGCGCAACGAAAAACAGATGCC	720	

QY 901 GTCATTACTTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTACGTGAATCA 960
 Db 901 GTCATTACTTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTACGTGAATCA 960
 QY 961 CATTTTGGGGCAAAATGGGATGTGGATGCCGAGCGGAATCCTTGGCAATTTTGAACCGA 1020
 Db 961 CATTTTGGGGCAAAATGGGATGTGGATGCCGAGGAATCCTTGGCAATTTTGAACCGA 1020
 QY 1021 TCAACGTTTGGCAAGCAATGGGATAGATCTGAAGAATCTTAACCTTGGTTCATGGCCGTT 1080
 Db 1021 TCAACGTTTGGCAAGCAATGGGATAGATCTGAAGAATCTTAACCTTGGTTCATGGCCGTT 1080
 QY 1081 GCGATTTCAGTGGTGGGCGCATCTTCTTGGGGTTCGAGGCACTTTTGGAGCGTTGGCGCT 1140
 Db 1081 GCGATTTCAGTGGTGGGCGCATCTTCTTGGGGTTCGAGGCACTTTTGGAGCGTTGGCGCT 1140
 QY 1141 GGCACATCGTTTGTGTGGACTGTGATCTTTTCTTAAATGCAATTAATACAAAAAT 1200
 Db 1141 GGCACATCGTTTGTGTGGACTGTGATCTTTTCTTAAATGCAATTAATACAAAAAT 1200
 QY 1201 CCGCATCCCTCAATTGTAAGGGGATCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 Db 1201 CCGCATCCCTCAATTGTAAGGGGATCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 QY 1261 CAGACAGGCGCT 1271
 Db 1261 CAGACAGGCGCT 1271

RESULT 7
 AP005274 340000 bp DNA linear BCT 08-AUG-2002
 LOCUS 1/10.
 DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
 ACCESSION AP005274 BA000036
 VERSION AP005274.1 GI:21322764
 KEYWORDS Corynebacterium glutamicum ATCC 13032
 SOURCE Corynebacterium glutamicum ATCC 13032
 ORGANISM Corynebacteriaceae; Actinomycetales; Bacteria; Actinobacteria; Corynebacteriaceae; Corynebacterium.

REFERENCE
 1 Nakagawa, S.
 Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 340000)
 AUTHORS Nakagawa, S.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
 COMMENT This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. and Kitasato University.

FEATURES
 source
 1..340000
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 1..1575
 /gene="Cg10001"
 1..1575
 /gene="Cg10001"
 /note="PF00308:Bacterial dnaA protein
 TIGR00362:dnaA: chromosomal replication initiator protein
 dnaA"
 /codon_start=1
 /transl_table=11
 /product="ATPase involved in DNA replication initiation"
 /protein_id="BAB97394.1"
 /db_xref="GI:21322765"

gene
 CDS
 4766..5302
 /gene="Cg10005"
 4766..5302
 /gene="Cg10005"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB97398.1"
 /db_xref="GI:21322769"
 /translation="MTDPIEQAFERIRAEAMRNRSVPOLNKNDAFRPPAPKGVEX
 /translation="WSONSSILETRQNVADLTLSQADSGFDLPTRORAYLNT
 KPIAVGVAVLSTPNAKAVTENDJADLTTRLSRMGRSESLAYSEVEQIIPET
 PAQOEKXPDPAPVSISSNKAQYEVGREGASLSDSERHSPAPSEPHATADPE
 PELATPQIPREPAPHPNPNREVSLNPKYTESFVIQFPNRPANAAVAVASPAFEN
 PLFISGGSGGLKTHLHAGVNYAOELQGLAIKIVSSEEPNDINSRDROQETFR
 RYRNDILMVDIOFLAGKEGTOEFETFNALQADKOILSSDRPKQJLTLEDRL
 RTREPGCLITDIOCPDELTRIAIMKKAQGTGTVDRLELILASPESSIRELEGA
 IRVSAYSLLINQPIDKEMALYALRDILPEPDMETTPVIMETVAEYFELSDTLRGA
 GTRIVAVARQALATLCHELTDMSLPKIGDVGSKDHTVYAPRKIRQENTERTDY
 DEIQLTQLIKSRGN"
 /complement(1594..1920)
 /gene="Cg10002"
 complement(1594..1920)
 /gene="Cg10002"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB97395.1"
 /db_xref="GI:21322766"
 /translation="MIVDNFVYLMIDFKVRLSPVCADVSPAYPOFLGSCAQGKFRG
 ITPPTKLTLEQVCNYTAVNKVVDNFVSMATIGVSPQNPCTTPKLQKPSDEGSRRA
 ELNFIY"
 2292..3476
 /gene="Cg10003"
 2292..3476
 /gene="Cg10003"
 /EC_number="2.7.7.7"
 /note="PF00712:DNA polymerase III beta subunit, N-terminal
 domain
 PF02767:DNA polymerase III beta subunit, central domain
 PF02768:DNA polymerase III beta subunit, C-terminal domain
 TIGR00663:dnaI: DNA polymerase III, beta subunit"
 /codon_start=1
 /transl_table=11
 /product="DNA polymerase III beta subunit"
 /protein_id="BAB97396.1"
 /db_xref="GI:21322767"
 /translation="MESQNSFVRAREDELVTAVAVANLPPKTPQVPLRAMLITAD
 BELEAGYDVSTVRSLAEVSGRIAVAGKLSEITGSLPNKPVFRIDGSAFV
 TCGSSRFELPLIDDYPMPLPGLPAGTGSINAKLFTBEVSGVSAAGDSDS.FMLTGV
 SMEIYGNQINLALATDRFLALRTFMEANDPELVNKLIPATTLIDNARSLSGINDS
 IDIANGTDOVGRBLFGVHTDNRRTTRILDADPNIAPILPKETHIAISVEIAPLV
 DAIKRVSVIYARNQIVLHFSQVILTAGTAGHAESETT.PCAFTGKELTIAFVGY
 LNDGSLVFTSRVAFGFTPSRPALMIFPEPMSANENGITQTDTYFTYILMVRLL
 PG"
 3585..4769
 /gene="Cg10004"
 3585..4769
 /gene="Cg10004"
 /note="TIGR00611:recf: Recf protein"
 /codon_start=1
 /transl_table=11
 /product="Recombinational DNA repair ATPase"
 /protein_id="BAB97397.1"
 /db_xref="GI:21322768"
 /translation="MHIRSLERDYSRWPELVKDDPGTTFVIGNGRQKTNVIAIG
 YIAHLSHRVSDAPLNRHAENARVSVANVQGEIAAHILIKHPNAQASLNTRKY
 KIPRELGVAKTVLAPEDIALYKBPERRRYDDIITATRPRAAGTADYDKLKO
 RNALKTATIALRIGVGTGEGAAULSDTWTGRLKGLAENVAARFALNELGKIX
 EAYTTAPESRPAAVNYKTTIDQGLSQSEFPDAGIIEATLTLELAKQRETERSSS
 VGPBRDVIDLMGDOQPAKGFASGETWFSALRIAEFNLISDQTDPLILIDVFESE
 LDAGREKELVGAQVEQVLTAAVHDLPEMLKKVLPAQHTVTVQDNGTGRISLIDV
 DP"

QY 841 AAGGCTGCTAAATGACACTGATTTCTCCCTGATTTCTCCTTGTGTGCGATATGCGA 900
DB 27569 AAGGCTGCTAAATGACACTGATTTCTCCCTGATTTCTCCTTGTGTGCGATATGCGA 275628
QY 901 GTCATTTACTTTGGGCTCCGGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACGATGCA 960
DB 27629 GTCATTTACTTTGGGCTCCGGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACGATGCA 276688
QY 961 CAATTTGTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 1020
DB 27689 CAATTTGTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 27748
QY 1021 TCAAGCTTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 1080
DB 27749 TCAAGCTTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 27808
QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGCTGACGACACCTTGTGAGCGTTGCGCT 1140
DB 27809 GCGATTACAGTGTGGCGCATCTTCTTGGCGGCTGACGACACCTTGTGAGCGTTGCGCT 27868
QY 1141 GCGACATCTGTTTGTGTGACTGTGATCTTTTCTAACTGCTAAATTAACAAAT 1200
DB 27869 GCGACATCTGTTTGTGTGACTGTGATCTTTTCTAACTGCTAAATTAACAAAT 27928
QY 1201 CCGATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTAGAAAGGCTTAAG 1260
DB 27929 CCGATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTAGAAAGGCTTAAG 27988
QY 1261 CAGACAGCGCT 1271
DB 27989 CAGACAGCGCT 27999

RESULT 8
AX120085
LOCUS AX120085 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent EP1108790.
ACCESSION AX120085 AX114121
VERSION AX120085.1 GI:14036800
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tareishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
KYOMA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source location/Qualifiers
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3,309,400) split in 11, seq 1
0.000.001 0.349.980 349.980"

ORIGIN
Query Match 100.0%; Score 1271; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGATCATGATCTTGCTTCAATATTGCAATAGCTAGTGAAGTGGCCAAAC 60
DB 27629 GGGCGATCATGATCTTGCTTCAATATTGCAATAGCTAGTGAAGTGGCCAAAC 276788
QY 61 TGGCAACAACATACCCGCAATGTGTGATGTTAGTGTGCAAAAAACGCAAGAT 120
DB 276789 TGGCAACAACATACCCGCAATGTGTGATGTTAGTGTGCAAAAAACGCAAGAT 276848
QY 121 TCAATTAAGCTTGAAGTGTGCGCATTCAGGACGCTTGAACCAAGATGAATTAAGTTA 180

DB 276849 TCAATTAAGCTTGAAGTGTGCGCATTCAGGACGCTTGAACCAAGATGAATTAAGTTA 276908
QY 181 TGGGCGCTACGAATGCGGCAAGTCTAAACCTCCCTTGTGCGAGTTTGGGCAATGTA 240
DB 276909 TGGGCGCTACGAATGCGGCAAGTCTAAACCTCCCTTGTGCGAGTTTGGGCAATGTA 276968
QY 241 CCGATTTGGAATTTGGGCTTGTGCTCTTGTGTTAATTAACGCGCTACGAATGTTGGGCAAC 300
DB 276969 CCGATTTGGAATTTGGGCTTGTGCTCTTGTGTTAATTAACGCGCTACGAATGTTGGGCAAC 277028
QY 301 CCGATTTGGAATTTGGGCTTGTGCTCTTGTGTTAATTAACGCGCTACGAATGTTGGGCAAC 360
DB 277029 CCGATTTGGAATTTGGGCTTGTGCTCTTGTGTTAATTAACGCGCTACGAATGTTGGGCAAC 277088
QY 361 TGGGCGCGAGGCGGCTTGGGCGCATTCGCGCTACCAATGCTGTGTAACCTTCCGCA 420
DB 277089 TGGGCGCGAGGCGGCTTGGGCGCATTCGCGCTACCAATGCTGTGTAACCTTCCGCA 277148
QY 421 CGATTTCTAATGCGTTTCAATCCCGCTGATGTGTCATAAAACCCCATTTGCCGTTCTA 480
DB 277149 CGATTTCTAATGCGTTTCAATCCCGCTGATGTGTCATAAAACCCCATTTGCCGTTCTA 277208
QY 481 TTGCGTTTGGCGCTTATGACGAGACCTACGACATCTGCGGCGGCGGCGGCGGCTG 540
DB 277209 TTGCGTTTGGCGCTTATGACGAGACCTACGACATCTGCGGCGGCGGCGGCGGCTG 277268
QY 541 GTGCGGCGGCGGCTTATGACGAGACCTACGACATCTGCGGCGGCGGCGGCGGCTG 600
DB 277269 GTGCGGCGGCGGCTTATGACGAGACCTACGACATCTGCGGCGGCGGCGGCGGCTG 277328
QY 601 TCTACCGGAGTGGGCGATTCGACAGATTTCTTTTGAATTAAGGCTTGAATTCG 660
DB 277329 TCTACCGGAGTGGGCGATTCGACAGATTTCTTTTGAATTAAGGCTTGAATTCG 277388
QY 661 CTTTGTCTCTCTTTTGTGACGCTGATTTGATTCCTCCGCAAGAAAGGATTCG 720
DB 277389 CTTTGTCTCTCTTTTGTGACGCTGATTTGATTCCTCCGCAAGAAAGGATTCG 277448
QY 721 TTCTGTGCTGCTGCGAGTTTGAAGCTTCAACATGCTTGTGTAATTCAGGTGAGC 780
DB 277449 TTCTGTGCTGCTGCGAGTTTGAAGCTTCAACATGCTTGTGTAATTCAGGTGAGC 277508
QY 781 CCAATTTGGGCGGCTGCTGATCTTCTTGGGCTGTGACATCCGCTACTTCTTGGG 840
DB 277509 CCAATTTGGGCGGCTGCTGATCTTCTTGGGCTGTGACATCCGCTACTTCTTGGG 277568
QY 841 AAGGCTGCTAAATGACACTGATTTCTCCCTGATTTCTCCTTGTGTGCGATATGCGA 900
DB 27569 AAGGCTGCTAAATGACACTGATTTCTCCCTGATTTCTCCTTGTGTGCGATATGCGA 275628
QY 901 GTCATTTACTTTGGGCTCCGGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACGATGCA 960
DB 27629 GTCATTTACTTTGGGCTCCGGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACGATGCA 276688
QY 961 CAATTTGTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 1020
DB 27689 CAATTTGTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 27748
QY 1021 TCAAGCTTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 1080
DB 27749 TCAAGCTTGGGCAAAATGGGCAATGTGATGTCAGACAGAAATCCTTGCATTTTGAACGCA 27808
QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGCTGACGACACCTTGTGAGCGTTGCGCT 1140
DB 27809 GCGATTACAGTGTGGCGCATCTTCTTGGCGGCTGACGACACCTTGTGAGCGTTGCGCT 27868
QY 1141 GCGACATCTGTTTGTGTGACTGTGATCTTTTCTAACTGCTAAATTAACAAAT 1200
DB 27869 GCGACATCTGTTTGTGTGACTGTGATCTTTTCTAACTGCTAAATTAACAAAT 27928
QY 1201 CCGATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTAGAAAGGCTTAAG 1260
DB 27929 CCGATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTAGAAAGGCTTAAG 27988

QY 1261 CAGACAGCGCT 1271
 Db 277989 CAGACAGCGCT 277999

RESULT 9
 AF454053/c 2105 bp DNA linear BCT 27-JUN-2002

DEFINITION Corynebacterium glutamicum BrnF (brnF), BrnF (brnF), and Lrp-like
 regulator (Lrp) gene, complete cds.

ACCESSION AF454053
 AF454053.1 GI:21311379

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 Kernerkecht N., Sahm H., Yen M.R., Patek M., Sailer U.M.H. Jr. and
 Eggeling L.

Export of L-isoleucine from Corynebacterium glutamicum: a
 two-gene-encoded member of a new translocator family
 J. Bacteriol. 184 (14), 3947-3956 (2002)

JOURNAL
 MEDLINE
 PUBMED
 22077265
 12081967

REFERENCE
 Kernerkecht N., Eggeling L. and Sahm H.

AUTHORS
 TITLE
 JOURNAL
 Submitted (07-Nov-2001) IBT-1, Forschungszentrum, Leo-Brandt Str.,
 Juelich 52425, Germany

FEATURES
 source
 Location/Qualifiers

1. 2105
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 complement(185..711)
 /gene="brnF"
 complement(385..711)
 /note="Lrp-E transporter component 1; branched-chain amino
 acid export protein"
 /codon_start=1
 /transl_table=11
 /product="BrnF"
 /protein_id="AA046685.1"
 /db_xref="GI:21311380"
 /translation="MTDFSCILVAVCAVTFPALRAVPFLILKPLESQFVGKAM
 NMPAGILALIFSTERSNAIDKTLTFLGILAVATVVAHLGGRTLLSVGATLIVF
 GVNLFF"
 complement(708..1463)
 /gene="brnF"
 complement(708..1463)
 /note="Lrp-E transporter component 2; branched-chain amino
 acid export protein"
 /codon_start=1
 /transl_table=11
 /product="BrnF"
 /protein_id="AA046686.1"
 /db_xref="GI:21311381"
 /translation="MKTDLHSLLEVSFSKALBPDDKGYRRIEIAQGLKTSLAAL
 GMPFIAIGLAVLYGYEMWAPLFSGLIFAGSTEMVIAVAGAAPGALATLTL
 VNRHVFVAFSPFLHVKNPIDARFVFLIDEAVALIDAVATPAGSAMPILSMQAFH
 SVWFGALGVAAELIPEIKLEBALGSLFVTLTLDSCRTKKQIPSLIAGLSFTI
 ALVIRGQALFALLIFLGLTIRYFLKAKK"
 1538..1993
 /gene="Lrp"
 1538..1993
 /gene="Lrp"
 /note="Lrp"
 /codon_start=1
 /transl_table=11

ORIGIN
 Query Match 99.1%; Score 1260; DB 1; Length 2105;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 /product="Lrp-like regulator"
 /protein_id="AA046687.1"
 /db_xref="GI:21311382"
 /translation="MKTDLHSLLEVSFSKALBPDDKGYRRIEIAQGLKTSLAAL
 GMPFIAIGLAVLYGYEMWAPLFSGLIFAGSTEMVIAVAGAAPGALATLTL
 VNRHVFVAFSPFLHVKNPIDARFVFLIDEAVALIDAVATPAGSAMPILSMQAFH
 SVWFGALGVAAELIPEIKLEBALGSLFVTLTLDSCRTKKQIPSLIAGLSFTI
 ALVIRGQALFALLIFLGLTIRYFLKAKK"

QY 1 GCGGATCAATGGAATCTAGCTTCATATATTCGACAAATAGCTAGTTAGAGTGGCAAC 60
 Db 1563 GCGGATCAATGGAATCTAGCTTCATATATTCGACAAATAGCTAGTTAGAGTGGCAAC 1504
 QY 61 TGGCAAAATCTACCCGGAATTTGTGTATTTAGTGTGCAAAAACGACAGAGAT 120
 Db 1503 TGGCAAAATCTACCCGGAATTTGTGTATTTAGTGTGCAAAAACGACAGAGAT 1444
 QY 121 TCATTCAGGCTGAGGTGTGCGCATCCAGGAGGCTGGAAACGATGATTAAGGTTA 180
 Db 1443 TCATTCAGGCTGAGGTGTGCGCATCCAGGAGGCTGGAAACGATGATTAAGGTTA 1384
 QY 181 TCGGCGCTAGGAATCGCGGAAGGTCTAAAAACCTCGCTGCGAGGTTGGGCAATGA 240
 Db 1383 TCGGCGCTAGGAATCGCGGAAGGTCTAAAAACCTCGCTGCGAGGTTGGGCAATGA 1324
 QY 241 CCGGATTTGATTTGCGTTGCTGCTGTTATTCATACGGCTAGCAATGTGTGGCAGC 300
 Db 1323 CCGGATTTGATTTGCGTTGCTGCTGTTATTCATACGGCTAGCAATGTGTGGCAGC 1264
 QY 301 CCACATGTTTCCGCGCTGATTTTCCGCGGCTCACCGAAATCTGTCTATCGCCCTCGT 360
 Db 1263 CCACATGTTTCCGCGCTGATTTTCCGCGGCTCACCGAAATCTGTCTATCGCCCTCGT 1204
 QY 361 TGTGGGCGAGCGCCCTGGGCGCCATCGCGCTCACCAATTCGTGTGAATTCGCGCA 420
 Db 1203 TGTGGGCGAGCGCCCTGGGCGCCATCGCGCTCACCAATTCGTGTGAATTCGCGCA 1144
 QY 421 GATATTCATGCGTTTTCATTTCCCGCTGCAATGTGTGCAAAAACCCCATTCGCCGTTCTA 480
 Db 1143 GATATTCATGCGTTTTCATTTCCCGCTGCAATGTGTGCAAAAACCCCATTCGCCGTTCTA 1084
 QY 481 TTGGGTTTGGCGCTTATTCGACGAAGCTTACGAGTCACTGCGGCGACGCGCGAGCTG 540
 Db 1083 TTGGGTTTGGCGCTTATTCGACGAAGCTTACGAGTCACTGCGGCGACGCGCGAGCTG 1024
 QY 541 GTGGGCGTGGGACTTATTCGATGCAATAGCGTTTCACTCTGAGTATTCGGCGG 600
 Db 1023 GTGGGCGTGGGACTTATTCGATGCAATAGCGTTTCACTCTGAGTATTCGGCGG 964
 QY 601 TCTACCGGAGTGGCGATTCGACAGTGTATTCCTTTGAAATTAAGGCGCTCGAGTTGG 660
 Db 963 TCTACCGGAGTGGCGATTCGACAGTGTATTCCTTTGAAATTAAGGCGCTCGAGTTGG 904
 QY 661 CTTTGTCTCTCTTTGTACGCGTGAATTCCTGCGGACGCAAAAAGAGATCCC 720
 Db 903 CTTTGTCTCTCTTTGTACGCGTGAATTCCTGCGGACGCAAAAAGAGATCCC 844
 QY 721 TTTCTGTGTGTGCAAGTTTGAAGCTTCAATGCTCTGTGTGTAATTCAGGTCAGG 780
 Db 843 TTTCTGTGTGTGCAAGTTTGAAGCTTCAATGCTCTGTGTGTAATTCAGGTCAGG 764
 QY 781 CTTATTTGCGGCGTGTGATCTTTCTGGGTCTGTGACCATCGGATCTTCTTGGG 840
 Db 763 CTTATTTGCGGCGTGTGATCTTTCTGGGTCTGTGACCATCGGATCTTCTTGGG 724
 QY 841 AAAGGCTGCTAAATGACATGATTTTCTCTGATTTCTCTGTTGTGCGAGTATGCA 900
 Db 723 AAAGGCTGCTAAATGACATGATTTTCTCTGATTTCTCTGTTGTGCGAGTATGCA 664
 QY 901 GTCAATTAATTTTGGCTCCGCGCGGTTCCGTTCTTAATTCCTTAAGCCCTACGTAATCA 960

```

Db      663  GTCAATTACTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTTAAGCCCTACGTAATCA 604
Qy      961  CAATTTGGGGCAAAATGGCGATGCGATGCGACAGGAATCCCTTGGCATTTTACCGCA 1020
Db      603  CAATTTGGGGCAAAATGGCGATGCGATGCGACAGGAATCCCTTGGCATTTTACCGCA 544
Qy      1021  TCAACGTTTCGACGAATGCGATGATGTAAGACTTACCTTACCTTGGTCTGATTCGCGTT 1080
Db      543  TCAACGTTTCGACGAATGCGATGATGTAAGACTTACCTTACCTTGGTCTGATTCGCGTT 484
Qy      1081  GCGATTACAGTGTGTGGCGCATTTTCTTGGCGGTGCAAGCACTTGTGTAGGCGTCTG 1140
Db      483  GCGATTACAGTGTGTGGCGCATTTTCTTGGCGGTGCAAGCACTTGTGTAGGCGTCTG 424
Qy      1141  GGCACCATCGTTTTTGTGTAAGTGTGATCTTTTCTTAAACTGATTAATACAAAAT 1200
Db      423  GGCACCATCGTTTTTGTGTAAGTGTGATCTTTTCTTAAACTGATTAATACAAAAT 364
Qy      1201  CCGCATGCGCTTCATTTGAAGGGGATGCGGATTTTGAAGAACTAG-AAAAGCTTAA 1259
Db      363  CCGCATGCGCTTCATTTGAAGGGGATGCGGATTTTGAAGAACTAG-AAAAGCTTAA 304
Qy      1260  GCGAGCAGGCGT 1271
Db      303  GCGAGCAGGCGT 292

RESULT 10
LOCUS      AR31954 753 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6613545.
ACCESSION AR31954
VERSION AR31954.1 GI:40115725
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 753)
AUTHORS    Kemerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE      Nucleotide sequences coding for the export of branched chain amino
JOURNAL    Patent: US 6613545-A 2 02-SBP-2003;
FEATURES
SOURCE      1..753
            /organism="Unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      101  GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACCCCTG 160
Db      1  GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACCCCTG 60
Qy      161  GAACCATGATTAAGGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 220
Db      61  GAACCATGATTAAGGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
Qy      221  GCTGCAAGTTTGGCGATGTAACCGATGATGATGCGTTGTCTTGTGTTATTCATATC 280
Db      121  GCTGCAAGTTTGGCGATGTAACCGATGATGATGCGTTGTCTTGTGTTATTCATATC 180
Qy      281  GGCTACGAATGAGGCGACGCCCACTGTTTCGCGGCTGATTTTGGCGGCTCCACGGA 340
Db      181  GGCTACGAATGAGGCGACGCCCACTGTTTCGCGGCTGATTTTGGCGGCTCCACGGA 240
Qy      341  ATGCTGTATCGCCCTCGTTGTGTGGCGAGCGCCCTGCGCGCATCGCGCTCACACA 400
Db      241  ATGCTGTATCGCCCTCGTTGTGTGGCGAGCGCCCTGCGCGCATCGCGCTCACACA 300

```

```

Qy      401  TTGCTGTGAACCTCCGCCAGGATTCATGCGTTTTCATTCGCCCTGCATGTGTCAAA 460
Db      301  TTGCTGTGAACCTCCGCCAGGATTCATGCGTTTTCATTCGCCCTGCATGTGTCAAA 360
Qy      461  AACCCATGGCCGTTTTCATTCGTTTTCGCGCTTATGAGAAAGCCATACGAGTCACT 520
Db      361  AACCCATGGCCGTTTTCATTCGTTTTCGCGCTTATGAGAAAGCCATACGAGTCACT 420
Qy      521  GCGGCAAGCCGCGACGCTGTGCGGCGTGCAGATTATTCATGCAATGCAATAGGTTAC 580
Db      421  GCGGCAAGCCGCGACGCTGTGCGGCGTGCAGATTATTCATGCAATGCAATAGGTTAC 480
Qy      581  TCTTACTGAGTATTCGCGCGCTCTCACCGAGTGTGCGATGCGAGATGATCTTTTGA 640
Db      481  TCTTACTGAGTATTCGCGCGCTCTCACCGAGTGTGCGATGCGAGATGATCTTTTGA 540
Qy      641  ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTGTGACGCTGATTCGATTCCTG 700
Db      541  ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTGTGACGCTGATTCGATTCCTG 600
Qy      701  CGAACGAAAAAGCAATCCCTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCTT 760
Db      601  CGAACGAAAAAGCAATCCCTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCTT 660
Qy      761  GTGGTAATTCAGATCAGGCGCTATTGCGGCGCTGCTGATCTTGGGCTGTGAGC 820
Db      661  GTGGTAATTCAGATCAGGCGCTATTGCGGCGCTGCTGATCTTGGGCTGTGAGC 720
Qy      821  ATCCGTAATTCAGATCAGGCGCTATTGCGGCGCTGCTGATCTTGGGCTGTGAGC 883
Db      721  ATCCGTAATTCAGATCAGGCGCTATTGCGGCGCTGCTGATCTTGGGCTGTGAGC 753

RESULT 11
LOCUS      AX120373 753 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 289 from Patent EP1108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE      Corynebacterium glutamicum
ORGANISM    Corynebacterium glutamicum
REFERENCE   1
AUTHORS    Nakagawa,S., Mizoguchi,H., Ando,S., Hayaishi,M., Ochiai,K.,
TITLE      Novel polynucleotides
JOURNAL    Patent: EP 1108790-A 289 20-JUN-2001;
            KIOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
SOURCE      1..753
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      101  GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACCCCTG 160
Db      1  GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACCCCTG 60
Qy      161  GAACCATGATTAAGGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 220
Db      61  GAACCATGATTAAGGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
Qy      221  GCTGCAAGTTTGGCGATGTAACCGATGATGATGCGTTGTCTTGTGTTATTCATATC 280
Db      121  GCTGCAAGTTTGGCGATGTAACCGATGATGATGCGTTGTCTTGTGTTATTCATATC 180

```

QY 281 GGTACGATAGTGGGAGCCCACTGTTTCCGCGCTGATTTTCGGGGGCTCCACCGAA 340
 Db 181 GGCTACGATAGTGGGAGCCCACTGTTTCCGCGCTGATTTTCGGGGGCTCCACCGAA 240
 QY 341 ATGCTGATACGCTGCTGTTGTGGGCGGAGCCCTGGGGGCGCATCGGCTCCACCA 400
 Db 241 ATGCTGATACGCTGCTGTTGTGGGCGGAGCCCTGGGGGCGCATCGGCTCCACCA 300
 QY 401 TTGCTGATACGCTGCTGTTGTGGGCGGAGCTTTCATTCGCGCTGATGTCGAAA 460
 Db 301 TTGCTGATACGCTGCTGTTGTGGGCGGAGCTTTCATTCGCGCTGATGTCGAAA 360
 QY 461 AACCCATTGCCGCTTCTATTCGGTTTCGGGCTTATTCGAGCAAGCTTACGATCACT 520
 Db 361 AACCCATTGCCGCTTCTATTCGGTTTCGGGCTTATTCGAGCAAGCTTACGATCACT 420
 QY 521 GCGGCGAGGCGCGGAGCTGTCGCGGCGAGCTTATTCGAAATGCAATAGCGTTTAC 580
 Db 421 GCGGCGAGGCGCGGAGCTGTCGCGGCGAGCTTATTCGAAATGCAATAGCGTTTAC 480
 QY 581 TCCCTACTGGGATTTGGGCGGCTCTCAACGAGTGGCGATCGCAAGTTGATTCCTTTGAA 640
 Db 481 TCCCTACTGGGATTTGGGCGGCTCTCAACGAGTGGCGATCGCAAGTTGATTCCTTTGAA 540
 QY 641 ATTAAGGCGCTGAGTGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 700
 Db 541 ATTAAGGCGCTGAGTGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 600
 QY 701 CGAAGCAAAAGCAGATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 760
 Db 601 CGAAGCAAAAGCAGATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 660
 QY 761 GTGTAATTCAGGTCAGGCTTATTTGGGCGGCTGATGTTTGGGCTGTTGAC 820
 Db 661 GTGTAATTCAGGTCAGGCTTATTTGGGCGGCTGATGTTTGGGCTGTTGAC 720
 QY 821 ATCCGTAATTCCT 853
 Db 721 ATCCGTAATTCCT 753

RESULT 12
 LOCUS AX137710 753 bp DNA linear PART 30-MAY-2001
 DEFINITION Sequence 2 from Patent EP1096010.
 ACCESSION AX137710
 VERSION AX137710.1 GI:14273889
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1. Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 2 02-MAY-2001.
 JOURNAL Degussa AG (DE) / FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 location/Qualifiers
 1. 753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1. >753
 /note="unassigned protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CA039925.1"
 /db_xref="GI:14273890"
 /db_xref="REMBL:CA039925"
 /translation="MKTQBIHSSLEVSSEKALPEDDKCYRRIETAGQIKLSIAGL
 GMPDIFAGLIVIQYGEWMAAPLFSGLIFASTMLVIALVGAAPLGAIALTTLL

CDS

1. 753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1. >753
 /note="unassigned protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CA039925.1"
 /db_xref="GI:14273890"
 /db_xref="REMBL:CA039925"
 /translation="MKTQBIHSSLEVSSEKALPEDDKCYRRIETAGQIKLSIAGL
 GMPDIFAGLIVIQYGEWMAAPLFSGLIFASTMLVIALVGAAPLGAIALTTLL

ORIGIN

UNFRTVFAFSFPLHVVNKPILARPSVPLAIDEAVYATAPAGMSAWRLISMOIAFH
 STWVGSLTGVAIAELIPIETKGFALCS.FVITITLDSCTRKQIPBLLAGISFTI
 ALVPIPGALFALLIFGLITIRFFIKAK"

Query Match 59.2%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 3.3e-196;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTTCGCAATCCAGGACGCTTG 160
 Db 1 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTTCGCAATCCAGGACGCTTG 60
 QY 161 GAACGATGATTAAGGTTATCGGCGCTAGCAATTCGCAAGGTCTAAACCTCCCT 220
 Db 61 GAACGATGATTAAGGTTATCGGCGCTAGCAATTCGCAAGGTCTAAACCTCCCT 120
 QY 221 GCTCAGGTTTGGGCAATGACCGAATGGTATTCGTTGCTCTGTTATTCAATAC 280
 Db 121 GCTCAGGTTTGGGCAATGACCGAATGGTATTCGTTGCTCTGTTATTCAATAC 180
 QY 281 GGTACGAATGTTGGGCAAGCCCACTGTTTCGGGCTGATTTTCGGGCGCTCCACGAA 340
 Db 181 GGTACGAATGTTGGGCAAGCCCACTGTTTCGGGCTGATTTTCGGGCGCTCCACGAA 240
 QY 341 ATGCTGATACGCTGCTGTTGTGGGCGGAGCTTATTCGAGCAAGCTTACGATCACT 400
 Db 241 ATGCTGATACGCTGCTGTTGTGGGCGGAGCTTATTCGAGCAAGCTTACGATCACT 300
 QY 401 TTGCTGATACGCTGCTGTTGTGGGCGGAGCTTATTCGAGCAAGCTTACGATCACT 460
 Db 301 TTGCTGATACGCTGCTGTTGTGGGCGGAGCTTATTCGAGCAAGCTTACGATCACT 360
 QY 461 AACCCATTGCCGCTTCTATTCGGTTTCGGGCTTATTCGAGCAAGCTTACGATCACT 520
 Db 361 AACCCATTGCCGCTTCTATTCGGTTTCGGGCTTATTCGAGCAAGCTTACGATCACT 420
 QY 521 GCGGCGAGGCGCGGAGCTGTCGCGGCGAGCTTATTCGAAATGCAATAGCGTTTAC 580
 Db 421 GCGGCGAGGCGCGGAGCTGTCGCGGCGAGCTTATTCGAAATGCAATAGCGTTTAC 480
 QY 581 TCCCTACTGGGATTTGGGCGGCTCTCAACGAGTGGCGATCGCAAGTTGATTCCTTTGAA 640
 Db 481 TCCCTACTGGGATTTGGGCGGCTCTCAACGAGTGGCGATCGCAAGTTGATTCCTTTGAA 540
 QY 641 ATTAAGGCGCTGAGTGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 700
 Db 541 ATTAAGGCGCTGAGTGGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 600
 QY 701 CGAAGCAAAAGCAGATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 760
 Db 601 CGAAGCAAAAGCAGATCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 660
 QY 761 GTGTAATTCAGGTCAGGCTTATTTGGGCGGCTGATGTTTGGGCTGTTGAC 820
 Db 661 GTGTAATTCAGGTCAGGCTTATTTGGGCGGCTGATGTTTGGGCTGTTGAC 720
 QY 821 ATCCGTAATTCCT 853
 Db 721 ATCCGTAATTCCT 753

RESULT 13
 LOCUS BD014992 753 bp DNA linear PART 27-AUG-2002
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014992
 VERSION BD014992.1 GI:2255799
 KEYWORDS JP 2001169788-A/2.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.

```

REFERENCE 1 (bases 1 to 753)
AUTHORS Kemerukunehito,N., Salm,H., Eggering,L. and Pfeifferle,M.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
JOURNAL method of isolating the same and utilization thereof
DEUSSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH
OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/2
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PI 27-OCT-1999 DE 19951708.8
PI NICOLE KEMERUKUNEHITO,HERMANN SALM,LOTHAR EGGERING,WALTER PI
PFEIFFERLE
PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/08//
PC C12N15/09,C12R1:15),(C12N1/21,C12R1:15),(C12P13/06,C12R1:15),
PC C12N15/00,C12R1:15)
CC bnf
FH Key
FT CDS Location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196; Mismatches 0; Gaps 0;
Matches 753; Conservative 0; Indels 0;

101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 60
Qy 161 GAACCAAGATGAATGAATTCGCGGCTCAAGAAATCCGGCAAGGCTTAAACCTCCCT 220
Db 61 GAACCAAGATGAATGAATTCGCGGCTCAAGAAATCCGGCAAGGCTTAAACCTCCCT 120
Qy 221 GCTGAGGTTTGGGATGTACCCGATTTGATTTGGTGTCTCTTGTATTCATATAC 280
Db 121 GCTGAGGTTTGGGATGTACCCGATTTGATTTGGTGTCTCTTGTATTCATATAC 180
Qy 281 GGTACCAATGTTGGGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAA 340
Db 181 GGTACCAATGTTGGGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAA 240
Qy 341 ATGCTGATCATGCGCTGTGTGGGCGCAGGCGCCCTGGGCGCATCGCGCTCACACA 400
Db 241 ATGCTGATCATGCGCTGTGTGGGCGCAGGCGCCCTGGGCGCATCGCGCTCACACA 300
Qy 401 TTGCTGTGAATCTCCGCAAGATTCATATGGGTTTCAATCCGGCTGATGTGTCAAA 460
Db 301 TTGCTGTGAATCTCCGCAAGATTCATATGGGTTTCAATCCGGCTGATGTGTCAAA 360
Qy 461 AACCCATTGCGCGTTTCAATCCGGTTTCCGGCTTATTCAGCAAGCTTACAGTACT 520
Db 361 AACCCATTGCGCGTTTCAATCCGGTTTCCGGCTTATTCAGCAAGCTTACAGTACT 420
Qy 521 GCGGCGAGGCGCGAGGCTGTGGCGGCTGAGCTTATTCGAATGCAATAGCGTTTAC 580
Db 421 GCGGCGAGGCGCGAGGCTGTGGCGGCTGAGCTTATTCGAATGCAATAGCGTTTAC 480
Qy 581 TCCCTAGGGATTTGGGGGCTGTCAACGGAAGTGGGATTCGAGAGTTATTCCTTTGAA 640
Db 481 TCCCTAGGGATTTGGGGGCTGTCAACGGAAGTGGGATTCGAGAGTTATTCCTTTGAA 540
Qy 641 ATTAAAGGCTTGAAGTTCCTTGTCTCTCTTTGTCAGCTGATCTTGATTCCTGC 700
Db 541 ATTAAAGGCTTGAAGTTCCTTGTCTCTCTTTGTCAGCTGATCTTGATTCCTGC 600
Qy 701 CGAAGGAAATCCCTTCTCTGCTGTGAGGTTTGAAGTTTACATTCCTT 760

```

```

Db 601 CGACGAAAACGCAAGATTCCTTCTCTGCTGTGCGAGGTTTGAAGCTTACCATTCCTT 660
Qy 761 GTGTATTCAGATTCAGGAGCTTATTTGGGCGCTGTGATTCCTTGGGTCTGTAC 820
Db 661 GTGTATTCAGATTCAGGAGCTTATTTGGGCGCTGTGATTCCTTGGGTCTGTAC 720
Qy 821 ATCCGATTCCTTCTCTGAGAAAGGCTGTAA 853
Db 721 ATCCGATTCCTTCTCTGAGAAAGGCTGTAA 753

RESULT 14
BD162490
LOCUS Novel polynucleotide. 753 bp DNA linear PAT 17-JUN-2003
DEFINITION
ACCESSION BD162490.1 GI:27868248
VERSION JP 2002191370-A/289.
KEYWORDS
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 753)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
TITLE Yokoi,H., Tateishi,N., Senoo,A., Ikeda,K. and Ozaki,A.
JOURNAL Novel polynucleotide
PATENT: JP 2002191370-A 289 09-JUN-2002;
KYOMA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/289
PD 09-JUN-2002 JP 2000405096
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HAHUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
PC 04,C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
G01N33/56,
PC G01N33/56,G01N33/58,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
PC C12N15/00,
PC C12N5/00,C12N15/00
CC Novel polynucleotide
FH Key
FT source Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum".
source

FEATURES
source

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196; Mismatches 0; Gaps 0;
Matches 753; Conservative 0; Indels 0;

101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 60
Qy 161 GAACCAAGATGAATGAATTCGCGGCTCAAGAAATCCGGCAAGGCTTAAACCTCCCT 220
Db 61 GAACCAAGATGAATGAATTCGCGGCTCAAGAAATCCGGCAAGGCTTAAACCTCCCT 120
Qy 221 GCTGAGGTTTGGGATGTACCCGATTTGATTTGGTGTCTCTTGTATTCATATAC 280
Db 121 GCTGAGGTTTGGGATGTACCCGATTTGATTTGGTGTCTCTTGTATTCATATAC 180
Qy 281 GGTACCAATGTTGGGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAA 340

```

```

Db 181 GGCCTACGATGATGAGGAGCCCGCCAGCTGTTTCCGACGATTTTCGCGGCTCCACCCAA 240
Qy 341 ATGCTGATCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 400
Db 241 ATGCTGATCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 300
Qy 401 TTGCTGTGATCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 460
Db 301 TTGCTGTGATCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 360
Qy 461 AACCCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 520
Db 361 AACCCATGCGCCTGTTGTGGGCGAGCGCCCTGGGAGCCATGCGCTGACCA 420
Qy 521 GCGGCGAGCGCCCGAGCTGTCGCGGTCGACCTTATCTCAATGCAATAGCGTTTAC 580
Db 421 GCGGCGAGCGCCCGAGCTGTCGCGGTCGACCTTATCTCAATGCAATAGCGTTTAC 480
Qy 581 TCTCTAGTGGATATGCGCGCTGTCACCGAGTGGCGATGCGAGATTTCTTTTAA 640
Db 481 TCTCTAGTGGATATGCGCGCTGTCACCGAGTGGCGATGCGAGATTTCTTTTAA 540
Qy 641 ATTAAGGCGCTGAGATTCGCTGCTCTCTTTGTCACGCTGACTTGGATTCTGCG 700
Db 541 ATTAAGGCGCTGAGATTCGCTGCTCTCTTTGTCACGCTGACTTGGATTCTGCG 600
Qy 701 CGAAGCAAAAGCAATCCCTTCTGCTGCTGCTGCGAGGTTGAGCTTCAACATGCTT 760
Db 601 CGAAGCAAAAGCAATCCCTTCTGCTGCTGCTGCGAGGTTGAGCTTCAACATGCTT 660
Qy 761 GTGGTATTCAGATGAGCGCTATTTGGGCGCTGCTGATCTTCTGTTGAGC 820
Db 661 GTGGTATTCAGATGAGCGCTATTTGGGCGCTGCTGATCTTCTGTTGAGC 720
Qy 821 ATCCGATCTTCTTCTGCGAAAGGCTGCTAA 853
Db 721 ATCCGATCTTCTTCTGCGAAAGGCTGCTAA 753

```

```

RESULT 15
AX137083/c 397 bp DNA linear PAT 30-MAY-2001
LOCUS AX137083
DEFINITION Sequence 7 from Patent EP1090993.
ACCESSION AX137083
VERSION AX137083.1 GI:14273428
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Moeckel B., Pfeifferle W., Puehler A., Kalinowski J. and Bathe B.
Nucleotide sequences coding for the lrp gene
Patent: EP 1090993-A 7 11-APR-2001;
JOURNAL Degussa AG (DE)
FEATURES
source
1.397
Location/Qualifiers
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="lrp-Teil 1"

```

ORIGIN

```

Query Match 28.8%; Score 365.6; DB 6; Length 397;
Best Local Similarity 98.9%; Pred. No. 2e-89;
Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 9 AATGAAATAGCTTATATATGCAATAGCTAGTGAAGTGGCAAACTGGCAACA 68
Db 376 AATGAAATAGCTTATATATGCAATAGCTAGTGAAGTGGCAAACTGGCAACA 317
Qy 69 AAATAACCGGCAATTGTGATGATGATGATGCAAAAAAGCAAGAGATTCAATCAA 128

```

```

Db 316 AAATACCGGCAATTGTGATGATGATGATGCAAAAAAGCAAGAGATTCAATCAA 257
Qy 129 GCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188
Db 256 GCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 197
Qy 189 ACGAATGCGCAAGGCTTAAAACTCCCTGCTGCAAGGTTGGGCAATGACCGATTG 248
Db 196 ACGAATGCGCAAGGCTTAAAACTCCCTGCTGCAAGGTTGGGCAATGACCGATTG 137
Qy 249 GATTTGGGTTGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 308
Db 136 GATTTGGGTTGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 77
Qy 309 TTTCCGCGCTGATTTTTCGCGGCTGCAACCGAAATGCTGATGATGATGATGATGATGAT 368
Db 76 TTTCCGCGCTGATTTTTCGCGGCTGCAACCGAAATGCTGATGATGATGATGATGATGAT 17
Qy 369 CAGGCGCCCTGG 380
Db 16 CAGGCGCCCTGG 5

```

Search completed: April 19, 2004, 18:29:50
Job time : 5394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 591.9 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-1
Perfect score: 1271
Sequence: 1 ggcgcagcatcagatcagtag.....aggcttaagcagacagcgcct 1271

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_29Jan04.*

1: geneseq119808.*
2: geneseq119908.*
3: geneseq200008.*
4: geneseq200108.*
5: geneseq200208.*
6: geneseq200308.*
7: geneseq200408.*
8: geneseq200508.*
9: geneseq200608.*
10: geneseq200708.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1271	100.0	1271	AAH21109	Aah21109 C. glutam
2	1271	100.0	1271	AAH21112	Aah21112 C. glutam
3	1271	100.0	349980	AAH64966	Aah64966 C. glutam
4	753	59.2	753	AAH21110	Aah21110 C. glutam
5	753	59.2	753	AAH65254	Aah65254 C. glutam
6	553	43.5	564	ACA01969	ACA01969 C. glutam
7	365.6	28.8	397	AAH61693	Aah61693 C. glutam
8	365.6	28.8	778	AAH61695	Aah61695 C. glutam
9	324	25.5	324	AAH21111	Aah21111 C. glutam
10	324	25.5	324	AAH65255	Aah65255 C. glutam
11	324	25.5	324	ACA01970	ACA01970 C. glutam
12	185	14.6	715	AAH61688	Aah61688 C. glutam
13	177	13.9	177	ACA01968	ACA01968 C. glutam
14	127.4	10.0	711	AAH66873	Aah66873 C. glutam
15	127.4	10.0	711	ACA01294	ACA01294 C. glutam
16	127.4	10.0	309400	AAH65534	Aah65534 C. glutam
17	48.2	3.8	2000	ADA71938	Ada71938 Rice gene
18	48	3.8	10732	AAH10594	Aah10594 Gene enco
19	40.8	3.2	7645	AAH59610	Aah59610 Human gen
20	40.2	3.2	4947	AAH08627	Aah08627 TSPS17 po
21	39.8	3.1	768	ADC91639	ADC91639 E. faeciu
22	39.2	3.1	6741	AAH10595	Aah10595 Gene enco
23	38.2	3.0	946	AAH55683	Aah55683 Human bre

24	37.6	3.0	759	7	ACF69125	Acf69125 Photorhab
25	37.6	3.0	5059	2	AAH84332	Aah84332 Stealth v
26	37.6	3.0	110000	7	ACF67367_20	Continuation (21 o
27	37.6	3.0	243072	7	ACF65382	ACF65382 Photorhab
28	37.2	2.9	1024	6	AAH66366	Aah66366 Helicobac
29	37.2	2.9	2000	7	ADA71938	Ada71938 Rice gene
30	36.8	2.9	1230	2	ACA42205	ACA42205 Proxayot
31	36.4	2.9	751	7	AAV58489	Aav58489 3' fragme
32	36.4	2.9	751	2	AAH61146	Aah61146 3' CDNA s
33	36.4	2.9	751	3	AAA06252	AAA06252 Human imm
34	36.4	2.9	751	3	AB571157	Ab571157 Human pro
35	36.4	2.9	751	4	AAH93368	Aah93368 Human pro
36	36.4	2.9	751	4	AAH63460	Aah63460 Human pro
37	36.4	2.9	751	4	AAH02433	Aah02433 Prostate
38	36.4	2.9	751	4	AAH84682	Aah84682 Human pro
39	36.4	2.9	751	5	ACA59269	ACA59269 Prostate
40	36.4	2.9	751	5	AAH10011	Aah10011 Human pro
41	36.4	2.9	751	6	AB194832	Ab194832 Human NI-
42	36.4	2.9	751	6	AB585841	Ab585841 Prostate
43	36.4	2.9	751	7	ACC94996	Acc94996 Prostate
44	36.4	2.9	751	9	ADB13462	Adb13462 Human pro
45	36.2	2.8	19012	2	AAQ36952	Aaq36952 HSA gene.

ALIGNMENTS

RESULT 1
ID AAH21109 standard; DNA; 1271 BP.

AC AAH21109;
DT 05-SEP-2001 (first entry)
XX C. glutamicum DNA encoding brnF and brnE.

DE L-amino acid production; brnF; brnE; branched-chain amino acid;
KW Corynebacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.

XX Corynebacterium glutamicum.

OS Key Location/Qualifiers

FT CDS 101..856

FT /tag= a "brnF"

FT /product= "b" 853..1179

FT /tag= b

FT /product= "brnE"

XX EPI096010-A1.

PN 02-MAY-2001.

PD 11-OCT-2000; 2000EP-00122057.

PF 27-OCT-1999; 99DE-01051708.

PR (DEGS) DEGUSSA AG.

PA (KERJ) FORSCHUNGSZENTRUM UEBELICH GMBH.

XX Kennernknecht N, Eggeling L, Sahn H, Pfefferle W;

PI WPI; 2001-391595/42.

DR P-PADB; AAB66247; AAB66248.

XX New export genes from coryneform bacteri, useful for increasing

PT fermentative production of branched-chain amino acids.

PS Claim 4 (1); Page 13; 23pp; German.

XX This invention describes a novel isolated polynucleotide (I) containing

XX This invention describes a novel isolated polynucleotide (I) containing

at least one sequence that (i) is 70% identical with a sequence that encodes at least one of 251 amino acid (aa) (3) or 108 aa (5) polypeptides; (ii) encodes a polypeptide at least 70% identical with (3) or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least 15 consecutive bases from (i)-(iii). The invention also describes (a) protein derived from sequences of 753 bp (2) or 324 bp (4); (b) coryneform microorganisms, especially *Corynebacterium*, transformed with one or more (i), where these are replicative DNA; (c) production of branched-chain L-2a by fermentation of coryneform bacteria in which the bnf and/or bnf genes (or equivalent sequences) are amplified, especially overexpressed; and (d) method for isolating the bnf and/or bnf genes. (i) is used for transformation of coryneform bacteria being used for fermentative production of branched-chain amino acids, specifically leucine, isoleucine and valine, which are useful in medicine and animal nutrition. (i) can also be used as source of primers and probes for isolation of related sequences. Transformation with (i) increases yield of branched-chain amino acids. This sequence encodes the *Corynebacterium glutamicum* ATCC 14752 bnf and bnf proteins described in the method of the invention

Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGGATCATGATGATCTAGCTTCATATATGCACAATAGCTTGAAGTGGCAAC 60
DB 1 GCGGATCATGATGATCTAGCTTCATATATGCACAATAGCTTGAAGTGGCAAC 60
QY 61 TGGCAAAACAACTACCCGGCAATGTGTGATGATGATGATGATGATGATGATGAT 120
DB 61 TGGCAAAACAACTACCCGGCAATGTGTGATGATGATGATGATGATGATGATGAT 120
QY 61 TGGCAAAACAACTACCCGGCAATGTGTGATGATGATGATGATGATGATGATGAT 120
DB 61 TGGCAAAACAACTACCCGGCAATGTGTGATGATGATGATGATGATGATGATGAT 120
QY 121 TCATCAAGCTCGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 121 TCATCAAGCTCGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGGTTGGGCAAT 240
DB 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGGTTGGGCAAT 240
QY 241 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 TGTGGGCGAGGCGCCCTGGGCGCATCGGCTCAACACATGCTGTGATGATGATGAT 420
DB 361 TGTGGGCGAGGCGCCCTGGGCGCATCGGCTCAACACATGCTGTGATGATGATGAT 420
QY 421 CGTATTCATGCGTTTCAATTCGCGTGCATGTGTGATGATGATGATGATGATGAT 480
DB 421 CGTATTCATGCGTTTCAATTCGCGTGCATGTGTGATGATGATGATGATGATGAT 480
QY 481 TTCGATTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGGCTG 540
DB 481 TTCGATTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGGCTG 540
QY 541 GTCGCGGTGGCGATTCATTCGAAATAGCGTTTCACTCTTACTGAGGATTTGGCG 600
DB 541 GTCGCGGTGGCGATTCATTCGAAATAGCGTTTCACTCTTACTGAGGATTTGGCG 600
QY 601 TCTCACCGAGTGGCGATTCGAAATAGCGTTTCACTCTTACTGAGGATTTGGCG 660
DB 601 TCTCACCGAGTGGCGATTCGAAATAGCGTTTCACTCTTACTGAGGATTTGGCG 660
QY 661 CTTTGTCTCTCTTTGTCAAGCTGATTTGATTTCTGCGGAGCAAAAAGAGATGCC 720
DB 661 CTTTGTCTCTCTTTGTCAAGCTGATTTGATTTCTGCGGAGCAAAAAGAGATGCC 720

```

```

QY 721 TTCTGTCTCTCTCTTTGTCAAGCTGATTTGATTTCTGCGGAGCAAAAAGAGATGCC 780
DB 721 TTCTGTCTCTCTCTTTGTCAAGCTGATTTGATTTCTGCGGAGCAAAAAGAGATGCC 780
QY 781 CCTATTGGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CCTATTGGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 AAAGCTCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGCGA 900
DB 841 AAAGCTCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGCGA 900
QY 901 GTCATTTACTTTGGCGCTCCGGCGGCTTCGTTTCTTAACTTAACTTAACTTAACT 960
DB 901 GTCATTTACTTTGGCGCTCCGGCGGCTTCGTTTCTTAACTTAACTTAACTTAACT 960
QY 961 CAATTTGTGGGCAAAATGCGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 CAATTTGTGGGCAAAATGCGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 TCAAGCTTTCGCGAGCAATGCGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 TCAAGCTTTCGCGAGCAATGCGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 GCGATTCAGTGTGTGCGGCTCTTCTGCGGCTGCAACGCACTTGTGAGGCTGGCT 1140
DB 1081 GCGATTCAGTGTGTGCGGCTCTTCTGCGGCTGCAACGCACTTGTGAGGCTGGCT 1140
QY 1141 GCGACCATGCTTTTGTGAGCTGTAATCTTTCTTAACTTAACTTAACTTAACT 1200
DB 1141 GCGACCATGCTTTTGTGAGCTGTAATCTTTCTTAACTTAACTTAACTTAACT 1200
QY 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAACTTAACT 1260
DB 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAACTTAACT 1260
QY 1261 CAGACAGCGCT 1271
DB 1261 CAGACAGCGCT 1271

```

RESULT 2

AAH21112
ID AAH21112 standard; DNA; 1271 BP.

AAH21112;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding bnf and bnfE.

L-amino acid production; bnf; bnfE; branched-chain amino acid; coryneform bacterium; leucine; isoleucine; valine; medicine; animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers
FH 101..856
FT /tag= "a"
FT /product= "bnf"
FT 853..1179
FT /tag= "b"
FT /product= "bnfE"

EP1096010-A1.

02-MAY-2001.

11-OCT-2000; 2000EP-00122057.

27-OCT-1999; 99DE-01051708.

XX 02-MAY-2001.
 XX 11-OCT-2000; 2000EP-00122057.
 XX 27-OCT-1999; 99DE-01051708.
 PR (DEGS) DEGUSSA AG.
 PA (KER) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX Kennerknecht N, Eggeling L, Salm H, Pfeffele W;
 XX WPI: 2001-391595/42.
 DR P-PSDB; AAB86247.
 XX New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 5; Page 14-15; 23pp; German.
 XX This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (1), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnB and/or brnF genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnB and/or
 CC brnF genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brnF protein described in the
 CC method of the invention
 XX
 XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 SQ
 Query Match 59.2%; Score 753; DB 4; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1.3e-230;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 361 AACCCATTGCCCGTTTCATTGCGTTTCGGCTTATCGAAGAGCTACAGACTCACT 420
 QY 521 GCGGCCAGGCGCCGACAGGCTGTGCGCTGGCGACTTATCTCAATGCAATAGCGTTTAC 580
 DB 421 GCGGCCAGGCGCCGACAGGCTGTGCGCTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 QY 581 TCGTACTGGGTATTCGCGCGGTCTCACCGAGTGGCGATCGACAGTGTATTCCTTTGAA 640
 DB 481 TCGTACTGGGTATTCGCGCGGTCTCACCGAGTGGCGATCGACAGTGTATTCCTTTGAA 540
 QY 641 ATTAGGCGCTCGAGTTCGCCCTTTGCTCTCTTTGTGACGCTGACTTGGATTCTTCG 700
 DB 541 ATTAGGCGCTCGAGTTCGCCCTTTGCTCTCTTTGTGACGCTGACTTGGATTCTTCG 600
 QY 701 CGAAGCAAAAAGAGATCCCTTCTGCGTGTGCGAGTTTGAAGTTCACATTGCTCTT 760
 DB 601 CGAAGCAAAAAGAGATCCCTTCTGCGTGTGCGAGTTTGAAGTTCACATTGCTCTT 660
 QY 761 GTGGTAATTCAGAGTCCGCTTATTGGCGCGCTGTGATCTTCTTGGGCTCTGTGACC 820
 DB 661 GTGGTAATTCAGAGTCCGCTTATTGGCGCGCTGTGATCTTCTTGGGCTCTGTGACC 720
 QY 821 ATCCGTAATCTTCTTGGGAAAGGCTGTAA 853
 DB 721 ATCCGTAATCTTCTTGGGAAAGGCTGTAA 753

RESULT 5
 ID AAB65254 standard; DNA; 753 BP.
 XX AAB65254;
 XX 26-SEP-2001 (first entry)
 DT C glutamicum coding sequence fragment SEQ ID NO: 289.
 XX
 DE Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 XX Corynebacterium glutamicum.
 OS
 XX EPI108790-A2.
 XX
 XX 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-00127688.
 XX
 XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tareishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI: 2001-376931/40.
 DR P-PSDB; AAG90035.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 289; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino

QY 769 TCCAGGTCAAGCCCTATTGCGGCGTTCGATCTTGTGGGTCTGTGACCATCCGGA 828
 DB 480 TCCAGGTCAAGCCCTATTGCGGCGTTCGATCTTGTGGGTCTGTGACCATCCGGA 539
 QY 829 CTTCTTCTGGGAAAGGCTGCTAA 853
 DB 540 CTTCTTCTGGGAAAGGCTGCTAA 564

RESULT 7

AAAF693/c
 ID AAAF693 standard; DNA; 397 BP.

AC AAAF693;
 XX

DT 12-JUL-2001 (first entry)
 XX

DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.
 XX

KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 XX medicine; animal feed supplement; de.
 XX

OS Corynebacterium glutamicum.
 XX

PN EP1090993-A1.
 XX

XX 11-APR-2001.
 XX

XX 29-SEP-2000; 2000EP-00121159.
 XX

XX 05-OCT-1999; 99DE-01047792.
 XX

PA (DEGS) DEGUSA-HUELS AG.
 XX

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
 XX

DR WPI; 2001-292927/31.
 XX

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 XX increased synthesis of amino acids, particularly lysine and isoleucine.
 XX

PS Example 3; Page 16; 22pp; German.
 XX

CC This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention
 CC

XX Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

Query Match 28.8%; Score 365.6; DB 4; Length 397;
 Best Local Similarity 98.9%; Pred. No. 2.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGAATCTAGCTCATATATATGACAAATAGCTAGTGGTGGCGAAACTGCAACA 68
 DB 376 AATGAATCTAGCTCATATATGACAAATAGCTAGTGGTGGCGAAACTGCAACA 317

QY 69 AAATACCCGCGCAATGTGTGATGATGTAGTGCAGAAAAACGCAAGATTCATTCAA 128
 DB 316 AAATACCCGCGCAATGTGTGATGATGTAGTGCAGAAAAACGCAAGATTCATTCAA 257

QY 129 GCGTGAAGTGTGCGCATCAAGGACGCGTGAACGATGATTAAGATTATCGGCGCT 188
 DB 129 GCGTGAAGTGTGCGCATCAAGGACGCGTGAACGATGATTAAGATTATCGGCGCT 188

DB 256 GCCTGAGAGTGTGCGCATCAAGGACGCGTGAACGATGATTAAGATTATCGGCGCT 197
 QY 189 AGGAATGCGCGAAGTCTTAAAAACCTCCCTTGCTGCGAGGTTTGGGCAATGACCGATTG 248
 DB 196 AGGAATGCGCGAAGTCTTAAAAACCTCCCTTGCTGCGAGGTTTGGGCAATGACCGATTG 137
 QY 249 GATTGCGTTGTGTCTCTGTGTTATTCAATACGCTACGAATGTGTGGCAGCCCACTGT 308
 DB 136 GATTGCGTTGTGTCTCTGTGTTATTCAATACGCTACGAATGTGTGGCAGCCCACTGT 77
 QY 309 TTTCGCGCTGATTTTTCGCGGAGCTCCACCGAAATCTGTGATGCGCTCTGTGTGGCG 368
 DB 76 TTTCGCGCTGATTTTTCGCGGAGCTCCACCGAAATGCTGTATGCGCTCTGTGTGGCG 17
 QY 369 CAGCGCCCTCG 380
 DB 16 CAGCGCCCTCG 5

RESULT 8

AAAF695/c
 ID AAAF695 standard; DNA; 778 BP.

AC AAAF695;
 XX

DT 12-JUL-2001 (first entry)
 XX

DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.
 XX

KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 XX medicine; animal feed supplement; ds.
 XX

OS Corynebacterium glutamicum.
 XX

PN EP1090993-A1.
 XX

XX 11-APR-2001.
 XX

XX 29-SEP-2000; 2000EP-00121159.
 XX

XX 05-OCT-1999; 99DE-01047792.
 XX

PA (DEGS) DEGUSA-HUELS AG.
 XX

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
 XX

DR WPI; 2001-292927/31.
 XX

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 XX increased synthesis of amino acids, particularly lysine and isoleucine.
 XX

PS Example 3; Page 17; 22pp; German.
 XX

CC This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention
 CC

XX Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;

Query Match 28.8%; Score 365.6; DB 4; Length 778;
 Best Local Similarity 98.9%; Pred. No. 3.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGAATCTAGCTCATATATATGACAAATAGCTAGTGGTGGCGAAACTGCAACA 68
 DB 9 AATGAATCTAGCTCATATATATGACAAATAGCTAGTGGTGGCGAAACTGCAACA 68

```

Db 376 AATGAATCTAGCTCATATATGCAATAGCCTAGTGGTGGCAAACTGGCAACA 317
QY 69 AAACATACCCGGGAATGTGTGTATGTATGTAGTGCAAAAAACCAAGATTCATTCAA 128
Db 316 AACTACCCCGGCAATGTGTGTATGTATGTATGTATGTATGTATGTATGTATGTAT 257
QY 129 GCTTGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 188
Db 256 GCTTGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 197
QY 169 ACGAATATGCGGCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248
Db 196 ACGAATATGCGGCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 137
QY 249 GATATGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308
Db 136 GATATGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 77
QY 309 TTTCCGCGCTGATTTTTCGCGGCTCCACGGAATGCTGTATGCGCTTGTGTGTGT 368
Db 76 TTTCCGCGCTGATTTTTCGCGGCTCCACGGAATGCTGTATGCGCTTGTGTGTGT 17
QY 369 CAGCGCCCTG 380
Db 16 CAGCGCCCTG 5

RESULT 9
AAH21111
ID AAH21111 standard; DNA; 324 BP.
AC AAH21111;
DT 05-SEP-2001 (first entry)
DE C. glutamicum b7m DNA.
XX
XX L-amino acid production; b7m; b7m; branched-chain amino acid;
KM coryneform bacterium; leucine; isoleucine; valine; medicine;
XX animal nutrition; ds.
XX
XX Corynebacterium glutamicum.
OS
XX EP1096010-A1.
XX
XX 02-MAY-2001.
XX
XX 11-OCT-2000; 2000EP-00122057.
XX
XX 27-OCT-1999; 99DE-01051708.
XX
XX (DEGS) DEGUSA AG.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Kennerknecht N, Eggeling L, Sahn H, Pfeifferle W;
XX
XX WPI; 2001-391595/42.
XX
XX P-PSDB; AAB86248.
XX
XX New export genes from coryneform bacteria, useful for increasing
PT fermentative production of branched-chain amino acids.
XX
XX Claim 5; Page 16; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (I) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with

```

```

CC one or more (I), where these are replicative DNA; (c) production of
CC branched-chain L-a by fermentation of coryneform bacteria in which the
CC b7m and/or b7m genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the b7m and/or
CC b7m genes. (I) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (I) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (I)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 14752 b7m protein described in the
CC method of the invention
XX
XX Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
SQ

Query Match 25.5%; Score 324; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.4e-93;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 853 ATGACAACTGATTTCTCCGTGATTTCCCTGTTGTGTGTGTGTGTGTGTGTGTGTGT 912
Db 1 ATGACAACTGATTTCTCCGTGATTTCTCCGTGTTGTGTGTGTGTGTGTGTGTGTGTGT 60
QY 913 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAAGTGAATCAATTTGTGGC 972
Db 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAAGTGAATCAATTTGTGGC 120
QY 973 AAAATGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
Db 121 AAAATGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 1033 AGCAATGCGATGATCTGAAGAAGCTTAACCTTGTGTGTGTGTGTGTGTGTGTGTGT 1092
Db 181 AGCAATGCGATGATCTGAAGAAGCTTAACCTTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 1093 GTGGCGCATTTTGTGGCGGTGACGCGACCTTGTGTGTGTGTGTGTGTGTGTGTGT 1152
Db 241 GTGGCGCATTTTGTGGCGGTGACGCGACCTTGTGTGTGTGTGTGTGTGTGTGTGT 300
QY 1153 TTTGTGGAATGATGATCTTTTC 1176
Db 301 TTTGTGGAATGATGATCTTTTC 324

RESULT 10
AAH65255
ID AAH65255 standard; DNA; 324 BP.
AC AAH65255;
DT 26-SEP-2001 (first entry)
DE C glutamicum coding sequence fragment SEQ ID NO: 250.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
XX EP108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Semoh A, Ikeda M, Ozaki A;
PI

```


XX WPI; 2001-376931/40.
 DR P-PSDB; AAG90036.
 XX
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 XX Claim 8; SEQ ID NO 290; 246bp + Sequence Listing; English.
 PS
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of Corynebacterium glutamicum, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium glutamicum. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 CC
 XX Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 25.5%; Score 324; DB 5; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.4e-93;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 853 ATGACAACTGATTTCTCTGTAATCTCCTTGTGTGCGAGTATGACATTAATCTTT 912
 Db 1 ATGACAACTGATTTCTCTGTAATCTCCTTGTGTGCGAGTATGAGATTAATCTTT 60
 QY 913 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGCATCAATTTGTGGC 972
 Db 61 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGCATCAATTTGTGGC 120
 QY 973 AAAATGGCATGTGATGCGACAGAGATCCTTGGCATTTTGACGCGATCAACGTTTGC 1032
 Db 121 AAAATGGCATGTGATGCGACAGAGATCCTTGGCATTTTGACGCGATCAACGTTTGC 180
 QY 1033 AGCAATGCGATGATCTGAAGACTTAACCTTTGTGATTTGCCGTTGCAATTAAGTG 1092
 Db 181 AGCAATGCGATGATCTGAAGACTTAACCTTTGTGATTTGCCGTTGCAATTAAGTG 240
 QY 1093 GTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTTGGCGGTGACCATCGTT 1152
 Db 241 GTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTTGGCGGTGACCATCGTT 300
 QY 1153 TTGTGTGACTGTGATCTTTTC 1176
 Db 301 TTGTGTGACTGTGATCTTTTC 324

RESULT 11

ACAO1970
 ID ACA01970 standard; DNA; 324 BP.
 XX ACA01970;
 AC
 XX 04-JUN-2003 (first entry)
 DT
 XX C. glutamicum derived ORF SEQ ID 1961.
 DE
 XX Coryneform; nucleic acid array; fermentation; culture; ds.
 KM
 XX Corynebacterium glutamicum.
 OS
 XX DE10126510-A1.
 PN
 XX 19-DEC-2002.
 PD
 XX 13-JUN-2001; 2001DE-01026510.
 PF

XX 13-JUN-2001; 2001DE-01026510.
 PR
 XX (DEGS) DEGUSA AG.
 PA
 XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Hutmacher K;
 XX WPI; 2003-279970/28.
 DR
 XX
 XX New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 XX
 XX Claim 1; Page 647; 709pp; German.
 PS
 XX This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA0010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention
 CC
 XX Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 25.5%; Score 324; DB 7; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.4e-93;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 853 ATGACAACTGATTTCTCTGTAATCTCCTTGTGTGCGAGTATGACATTAATCTTT 912
 Db 1 ATGACAACTGATTTCTCTGTAATCTCCTTGTGTGCGAGTATGAGATTAATCTTT 60
 QY 913 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGCATCAATTTGTGGC 972
 Db 61 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGCATCAATTTGTGGC 120
 QY 973 AAAATGGCATGTGATGCGACAGAGATCCTTGGCATTTTGACGCGATCAACGTTTGC 1032
 Db 121 AAAATGGCATGTGATGCGACAGAGATCCTTGGCATTTTGACGCGATCAACGTTTGC 180
 QY 1033 AGCAATGCGATGATCTGAAGACTTAACCTTTGTGATTTGCCGTTGCAATTAAGTG 1092
 Db 181 AGCAATGCGATGATCTGAAGACTTAACCTTTGTGATTTGCCGTTGCAATTAAGTG 240
 QY 1093 GTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTTGGCGGTGACCATCGTT 1152
 Db 241 GTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTTGGCGGTGACCATCGTT 300
 QY 1153 TTGTGTGACTGTGATCTTTTC 1176
 Db 301 TTGTGTGACTGTGATCTTTTC 324

RESULT 12

AAFe1688/c
 ID AAF61688 standard; DNA; 715 BP.
 XX AAF61688;
 AC
 XX 12-JUL-2001 (first entry)
 DT
 XX C. glutamicum lrp encoding DNA.
 DE
 XX Fermentation; L-lysine acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX
 XX Corynebacterium glutamicum.
 OS
 XX Key
 FH -35_signal Location/Qualifiers
 PF 62. .67


```
Db 233 TCATGATGATTTCCGCCGCAATTTCTAGGTCACCTTCCACGCCACCGCATCAAGT 292
Qy 462 ACCCATTTGCCGCTTTCTATTTGGTTTGGGCTTATGACGAGCTTACGAGTCACTG 521
Db 293 CCGGGCGCGCGCGCTTATTCACCTTACCGAGCTTACCGAGCTTACCGCATCGGT 352
Qy 522 CCGCCAGGCGCGCAGGCTGTGCGGTG--GCGACTTATCTCAATGCAATAGCGTTTC 578
Db 353 CAGCCCGCCCACTGCGCATATCAGTGGACGCGGGTGCTTACCGTTCAATTTGTGCC 412
Qy 579 ACTCTACTGGGTATTGCGGCTCTCACCGAGTGGCGCATGCGAGATTGATCTTTTG 638
Db 413 AAGCTCTGTGGGTATCCAGGAATATTGGCGCTTGTGTCAAGTGTGCCGATG 472
Qy 639 AAATTAAGGGCTCGAGTTCGCCCTTGTCTCTCTTGTCAAGCTGACTTGAATCCT 698
Db 473 ATCTAAAAGCATGATTTTGCCTGACCGCTGTGTGTGCTGGCGTGGAGGCAAT 532
Qy 699 GCCGAACGAAAAAGCAGATCCCTCTCTGTGCTGCTGCAAGTTGAGCTTCAACATTCCTC 758
Db 533 TCAAAAATACAGGATTTATCGTGCATTTATCGCGGTGATTTGCTGTGTTCCG 592
Qy 759 TTGTGTATTTCCAGTTCAGGCTTATTTGGCGGCTGTGATCTTCTTGGGTCTGTGA 818
Db 593 GTTTTGGGCGCCGAGCAGATGTGTTATGCTTTGACACGTACTTTTGAATCTTC 652
Qy 819 CCATCGGTACTTCTTC 835
Db 653 TTCTCGCGCTCGCTTC 669
```

Search completed: April 19, 2004, 11:41:56
Job time: 595.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 107.018 Seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271
Sequence: 1 gcgcgcacatgaatgaatctag.....aggctcaagcagacagcgc 1271

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: Issued Patents NA:*
- 2: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	US-09-471-803A-1	Sequence 1, Appli
2	1271	100.0	1271	US-09-471-803A-6	Sequence 6, Appli
3	753	59.2	753	US-09-471-803A-2	Sequence 2, Appli
4	324	25.5	324	US-09-471-803A-4	Sequence 4, Appli
5	58.6	4.6	756	US-09-489-039A-3016	Sequence 3016, Ap
6	32.4	4.1	7218	US-08-233-463-14	Sequence 14, Appl
7	39.8	3.1	768	US-09-107-532A-1266	Sequence 1266, Ap
8	39.2	3.1	1038	US-09-252-991A-11376	Sequence 11376, A
9	39.2	3.1	1086	US-09-252-991A-11418	Sequence 11418, A
10	39	3.1	399	US-09-621-976-8976	Sequence 8976, Ap
11	37	2.9	474	US-09-621-976-18033	Sequence 18033, A
12	36.8	2.9	1140	US-09-252-991A-3205	Sequence 3205, Ap
13	36.8	2.9	2061	US-09-252-991A-2761	Sequence 2761, Ap
14	36.4	2.9	751	US-09-020-956-12	Sequence 12, Appl
15	36.4	2.9	751	US-09-030-607-12	Sequence 12, Appl
16	36.4	2.9	751	US-09-438-313-12	Sequence 12, Appl
17	36.4	2.9	751	US-09-332-616A-12	Sequence 12, Appl
18	36.4	2.9	751	US-09-232-149A-12	Sequence 12, Appl
19	36.4	2.9	751	US-09-159-812-12	Sequence 12, Appl
20	36.4	2.9	751	US-09-636-215-12	Sequence 12, Appl
21	36.4	2.9	751	US-09-685-166A-12	Sequence 12, Appl
22	36.4	2.9	751	US-09-115-453-12	Sequence 12, Appl
23	36.4	2.9	751	US-09-688-489-12	Sequence 12, Appl
24	35.4	2.8	430	US-09-621-976-16656	Sequence 16656, A
25	35.2	2.8	801	US-09-328-352-2399	Sequence 2399, Ap
26	35.2	2.8	4403765	US-09-103-840A-2	Sequence 2, Appli
27	35.2	2.8	4411529	US-09-103-840A-1	Sequence 1, Appli

28	34.8	2.7	867	4	US-09-489-039A-5047	Sequence 5047, Ap
29	34.6	2.7	666	2	US-08-875-034A-1	Sequence 1, Appli
30	34.4	2.7	744	4	US-09-489-039A-1307	Sequence 1307, Ap
31	34.4	2.7	1224	4	US-09-266-965-22	Sequence 22, Appl
32	34.4	2.7	12349	4	US-09-266-965-74	Sequence 74, Appl
33	34.4	2.7	18331	4	US-09-266-965-96	Sequence 96, Appl
34	34.2	2.7	594	4	US-09-252-991A-13729	Sequence 13729, A
35	34.2	2.7	678	3	US-09-252-991A-13832	Sequence 13832, A
36	33.4	2.6	1875	3	US-09-422-869-21	Sequence 21, Appl
37	33.4	2.6	2949	3	US-08-433-522A-1	Sequence 1, Appli
38	33.4	2.6	2949	3	US-09-135-166-1	Sequence 1, Appli
39	33.4	2.6	2949	3	US-08-942-046-1	Sequence 1, Appli
40	33.4	2.6	2950	3	US-08-433-522A-5	Sequence 5, Appli
41	33.4	2.6	2950	3	US-09-135-166-5	Sequence 5, Appli
42	33.4	2.6	2950	3	US-08-942-046-5	Sequence 5, Appli
43	33.4	2.6	2984	3	US-08-433-522A-3	Sequence 3, Appli
44	33.4	2.6	2984	3	US-09-135-166-3	Sequence 3, Appli
45	33.4	2.6	2984	3	US-08-942-046-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1									
US-09-471-803A-1									
Sequence 1, Application US/09471803A									
Patent No. 6613545									
GENERAL INFORMATION:									
APPLICANT: KENNERKNECHT, NICOLE									
APPLICANT: SAHM, HERMANN									
APPLICANT: EGELING, LOTMAR									
APPLICANT: PEEPERLE, WALTER									
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF									
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE									
FILE REFERENCE: 21123/265496/MAS									
CURRENT APPLICATION NUMBER: US/09/471, 803A									
CURRENT FILING DATE: 1999-12-23									
PRIOR APPLICATION NUMBER: DE 199 51 708.8									
PRIOR FILING DATE: 1999-10-27									
NUMBER OF SEQ ID NOS: 12									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 1									
LENGTH: 1271									
TYPE: DNA									
ORGANISM: Corynebacterium glutamicum									
FEATURE:									
NAME/KEY: gene									
LOCATION: (101)..(853)									
OTHER INFORMATION: bmrF									
NAME/KEY: gene									
LOCATION: (853)..(1176)									
OTHER INFORMATION: bmrE									
OTHER INFORMATION: ATCC14752									
US-09-471-803A-1									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GGCGGATCAATGGAATCTGCTTCAATATTTGCACAAATGCGCTAGTGGTGGCGGAAC	60						
DB	1	GGCGGATCAATGGAATCTGCTTCAATATTTGCACAAATGCGCTAGTGGTGGCGGAAC	60						
QY	61	TGGCAACAAACTACCGCGCAATTGTGTGATGATTTAGTGGCAAAAACCAAGAGAT	120						
DB	61	TGGCAACAAACTACCGCGCAATTGTGTGATGATTTAGTGGCAAAAACCAAGAGAT	120						
QY	121	TCAATTAAAGCTGGAGAGTGTCCCATCCAAAGCAGCCCTGGAAACAGATGATTAAGGTTA	180						
DB	121	TCAATTAAAGCTGGAGAGTGTCCCATCCAAAGCAGCCCTGGAAACAGATGATTAAGGTTA	180						
QY	181	TCGGCCCTACGAATGCGCAAGGTCTTAAAACTCCCTGCTGGAGGTTTGGGCAATGTA	240						

```

Db      181  TCGGCGCTACGAAATCGCGCAGAGTCTAAAAACCTCCCTGCTGAGGTTTGCGCATGTA 240
Qy      241  CCCGATTTGATTTGCGTTTGGTCTCTTGGTTATTCATATACGAGCTACGAATGATGAGGAGC 300
Db      241  CCCGATTTGATTTGCGTTTGGTCTCTTGGTTATTCATATACGAGCTACGAATGATGAGGAGC 300
Qy      301  CCCGATTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGATTCGCTCTGT 360
Db      301  CCCGATTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGATTCGCTCTGT 360
Qy      361  TGTGGGCGGAGGCGCCCTGGGCGCATCGCGCTCACCATCTGCGTGAATCCCGCA 420
Db      361  TGTGGGCGGAGGCGCCCTGGGCGCATCGCGCTCACCATCTGCGTGAATCCCGCA 420
Qy      421  CGTATTCATGCGTTTTCATTCGCGTGCATGCTGCAAAAACCCCATTCGCGGTTCTA 480
Db      421  CGTATTCATGCGTTTTCATTCGCGTGCATGCTGCAAAAACCCCATTCGCGGTTCTA 480
Qy      481  TTGCGTTTTCGCGCTTATTCAGAGAACTACGAGTCACTGCGGCGAGCCCGAGGCTG 540
Db      481  TTGCGTTTTCGCGCTTATTCAGAGAACTACGAGTCACTGCGGCGAGCCCGAGGCTG 540
Qy      541  GTGCGGCTGAGCACTTATTCATATGCAATAGCGTTTCACTCTGAGGTTTTCGCGG 600
Db      541  GTGCGGCTGAGCACTTATTCATATGCAATAGCGTTTCACTCTGAGGTTTTCGCGG 600
Qy      601  TCTCACCGAGATGCGCATGCGAGAGTTATTCCTTTGAAATTAAGGCGCTGAGATTGC 660
Db      601  TCTCACCGAGATGCGCATGCGAGAGTTATTCCTTTGAAATTAAGGCGCTGAGATTGC 660
Qy      661  CCTTGTCTCTCTTTTGTGACGCTGACTTTGAGATTCCTGCGAGAAAGAAACAGATCC 720
Db      661  CCTTGTCTCTCTTTTGTGACGCTGACTTTGAGATTCCTGCGAGAAAGAAACAGATCC 720
Qy      721  TTCTGTCTGTCTGCGAGGTTTGAAGTTACCAATGCTTGTGTGATTCAGTCAAGC 780
Db      721  TTCTGTCTGTCTGCGAGGTTTGAAGTTACCAATGCTTGTGTGATTCAGTCAAGC 780
Qy      781  CCTATTGCGGCGCTGCTGATCTCTTGGGTCGTTGAGCAATCCGCTTCTTCTTGGG 840
Db      781  CCTATTGCGGCGCTGCTGATCTCTTGGGTCGTTGAGCAATCCGCTTCTTCTTGGG 840
Qy      841  AAAAGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGCA 900
Db      841  AAAAGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGCA 900
Qy      901  GTGATTTACTTTTGGGCTCCGCGGCTCCGTTCTTAAATCTTAAAGCCCTAGATGCA 960
Db      901  GTGATTTACTTTTGGGCTCCGCGGCTCCGTTCTTAAATCTTAAAGCCCTAGATGCA 960
Qy      961  CAATTTGTGGGCAAAATGCGAGATGAGATCCAGAGAAATCTTGGCAATTTGACCGCA 1020
Db      961  CAATTTGTGGGCAAAATGCGAGATGAGATCCAGAGAAATCTTGGCAATTTGACCGCA 1020
Qy      1021  TCAAGCTTTGCGAGCAATGCGAATGATCTGAAGACTTAACTTTGCTCATTCGCTT 1080
Db      1021  TCAAGCTTTGCGAGCAATGCGAATGATCTGAAGACTTAACTTTGCTCATTCGCTT 1080
Qy      1081  GCGATTTACAGTGTGCGCATCTTCTTGGGCTGACGACCTTGTGAGCGTTGGCGCT 1140
Db      1081  GCGATTTACAGTGTGCGCATCTTCTTGGGCTGACGACCTTGTGAGCGTTGGCGCT 1140
Qy      1141  GCGACCATGCTTTTGTGAGTGTGATCTTTTCTAAATCTGCTAAATAAACAATAAT 1200
Db      1141  GCGACCATGCTTTTGTGAGTGTGATCTTTTCTAAATCTGCTAAATAAACAATAAT 1200
Qy      1201  CCGCATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
Db      1201  CCGCATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
Qy      1261  CAGACAGCGCT 1271
Db      1261  CAGACAGCGCT 1271

```

```

Db      1261  CAGACAGCGCT 1271

RESULT 2
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERNKCHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGSELING, LOTHAR
; APPLICANT: PEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; OTHER INFORMATION: ATCC13032
; OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match      100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  GCGGATCATAGGAATGATCTGATCATATATGACAAATAGCTAGTGAAGTCCGAAAC 60
Db      1  GCGGATCATAGGAATGATCTGATCATATATGACAAATAGCTAGTGAAGTCCGAAAC 60
Qy      61  TGGCAACAAACTACCCGGCAATGTGTGATGATGTGTCACAAACGCAAGAGAT 120
Db      61  TGGCAACAAACTACCCGGCAATGTGTGATGATGTGTCACAAACGCAAGAGAT 120
Qy      121  TCATTCAAGCTGAGAGGTGTCGCATCCAGAGCAAGCCCTGGAAACGATGATAAGGTTA 180
Db      121  TCATTCAAGCTGAGAGGTGTCGCATCCAGAGCAAGCCCTGGAAACGATGATAAGGTTA 180
Qy      181  TCGGCGCTACGAAATCGCGCAGAGTCTAAACCTCCCTGCTGACAGTTTGGGCAATGTA 240
Db      181  TCGGCGCTACGAAATCGCGCAGAGTCTAAACCTCCCTGCTGACAGTTTGGGCAATGTA 240
Qy      241  CCCGATTTGATTTGCGTTTGGTCTCTTGGTTATTCATATACGAGCTACGAATGATGAGGAGC 300
Db      241  CCCGATTTGATTTGCGTTTGGTCTCTTGGTTATTCATATACGAGCTACGAATGATGAGGAGC 300
Qy      301  CCCGATTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGATTCGCTCTGT 360
Db      301  CCCGATTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGATTCGCTCTGT 360
Qy      361  TGTGGGCGGAGGCGCCCTGGGCGCATCGCGCTCACCATCTGCGTGAATCCCGCA 420
Db      361  TGTGGGCGGAGGCGCCCTGGGCGCATCGCGCTCACCATCTGCGTGAATCCCGCA 420
Qy      421  CGTATTCATGCGTTTTCATTCGCGTGCATGCTGCAAAAACCCCATTCGCGGTTCTA 480
Db      421  CGTATTCATGCGTTTTCATTCGCGTGCATGCTGCAAAAACCCCATTCGCGGTTCTA 480
Qy      481  TTGCGTTTTCGCGCTTATTCAGAGAACTACGAGTCACTGCGGCGAGCCCGAGGCTG 540
Db      481  TTGCGTTTTCGCGCTTATTCAGAGAACTACGAGTCACTGCGGCGAGCCCGAGGCTG 540

```

```

Db      481  TTGGTTTGGCGCTTTCAGCAGAACTTCAGCATCGCGCCAGCGCCGAGGCTG 540
Qy      541  GTGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCACTCTACTAGGTAATTCGGCG 600
Db      541  GTGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCACTCTACTAGGTAATTCGGCG 600
Qy      601  TCTCACCAGGAGTGGCGATCGCAGAGTTGATCTCTTTGAAATTAAAGGCTCGAGTTCC 660
Db      601  TCTCACCAGGAGTGGCGATCGCAGAGTTGATCTCTTTGAAATTAAAGGCTCGAGTTCC 660
Qy      661  CCTTGGCTCTCTCTTGTGACAGCTTGAAGTTTCTGCGGAAAGCAAAAGAGATCCC 720
Db      661  CCTTGGCTCTCTCTTGTGACAGCTTGAAGTTTCTGCGGAAAGCAAAAGAGATCCC 720
Qy      721  TTCTCTGCTGCTCGCAGGTTTGAAGCTTCAACCATTCCTTGTGTGTAATTCAGGTGAGC 780
Db      721  TTCTCTGCTGCTCGCAGGTTTGAAGCTTCAACCATTCCTTGTGTGTAATTCAGGTGAGC 780
Qy      781  CCTATTGCGGCGCTGCTGATCTTCTGCGGCTGTGACCATTCGGTAATTCCTTCTG 840
Db      781  CCTATTGCGGCGCTGCTGATCTTCTGCGGCTGTGACCATTCGGTAATTCCTTCTG 840
Qy      841  AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTCCTCTTGTGTGAGAGTAATGCA 900
Db      841  AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTCCTCTTGTGTGAGAGTAATGCA 900
Qy      901  GTCTATTCTTTGGCGCTTCGGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCA 960
Db      901  GTCTATTCTTTGGCGCTTCGGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCA 960
Qy      961  CAATTGTGGGCAAAATGGCGATGAGATGCGCAGAGGAATCTTGGCATTTTGAACCGCA 1020
Db      961  CAATTGTGGGCAAAATGGCGATGAGATGCGCAGAGGAATCTTGGCATTTTGAACCGCA 1020
Qy      1021  TCAAGCTTTCGACGAATGCAATGATCTGAAGACTTAACTTTCCTTGTCTATTCGCGTT 1080
Db      1021  TCAAGCTTTCGACGAATGCAATGATCTGAAGACTTAACTTTCCTTGTCTATTCGCGTT 1080
Qy      1081  GCGATTACAGTGGTGGCGCATCTTCTGCGGCTGACGACCTTGTGAGGCTTGGCGCT 1140
Db      1081  GCGATTACAGTGGTGGCGCATCTTCTGCGGCTGACGACCTTGTGAGGCTTGGCGCT 1140
Qy      1141  GCGACCATCGTTTGTGTGACTGTGATCTTTCTTAAACTGCAATAAATACAAAAT 1200
Db      1141  GCGACCATCGTTTGTGTGACTGTGATCTTTCTTAAACTGCAATAAATACAAAAT 1200
Qy      1201  CGGATGCGCTCAATTTGAAGGGGATGGGGATTTTAAAGAACTTAAAGGCTTAAG 1260
Db      1201  CGGATGCGCTCAATTTGAAGGGGATGGGGATTTTAAAGAACTTAAAGGCTTAAG 1260
Qy      1261  CAGACAGCGCT 1271
Db      1261  CAGACAGCGCT 1271

```

```

RESULT 3
US-09-471-803A-2
; Sequence 2, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTHAR
; APPLICANT: PEEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265486/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708-8
; PRIOR FILING DATE: 1999-10-27

```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: bmf
; OTHER INFORMATION: ATCC14752
; US-09-471-803A-2

```

```

Query Match      59.2% Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 6,66-243; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

```

```

Qy      101  GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGCGGCTG 160
Db      1  GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGCGGCTG 160
Qy      161  GAACCGATGATTAAGCTTATCGGCGCTACGAATGCGGCAAGTCTAAACCTCCCT 220
Db      61  GAACCGATGATTAAGCTTATCGGCGCTACGAATGCGGCAAGTCTAAACCTCCCT 120
Qy      221  GCTGCAAGTTGGGCAATGACCCGATGATGATGCTTGTCTCTTGTATTCATAC 280
Db      121  GCTGCAAGTTGGGCAATGACCCGATGATGATGATGCTTGTCTCTTGTATTCATAC 180
Qy      281  GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTATTTGCGGGGCTCCACGAA 340
Db      181  GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTATTTGCGGGGCTCCACGAA 240
Qy      341  ATGCTGATCAGCCCTCGTGTGAGGCGAGCGCCCTG3GCGCATCGCGCTACACACA 400
Db      241  ATGCTGATCAGCCCTCGTGTGAGGCGAGCGCCCTG3GCGCATCGCGCTACACACA 300
Qy      401  TTCTGATGAATCTTCGCGCAAGTATTCATTCGTTTCAATCCCGCTGATGTGCAAA 460
Db      301  TTCTGATGAATCTTCGCGCAAGTATTCATTCGTTTCAATCCCGCTGATGTGCAAA 360
Qy      461  AACCCATTGCGGCTTCTATTGATGATTTGCGGCTTATGACGAAGCTACGAGTCACT 520
Db      361  AACCCATTGCGGCTTCTATTGATGATTTGCGGCTTATGACGAAGCTACGAGTCACT 420
Qy      521  GCGGCGAGCGCCGAGGCTGTGTGCGGTGCGCACTTATCTCAATGCAATAGGCTTCA 580
Db      421  GCGGCGAGCGCCGAGGCTGTGTGCGGTGCGCACTTATCTCAATGCAATAGGCTTCA 480
Qy      581  TCTTACTGGTATTTGCGGCGTCTCACCGAGTGGGATGCGAGATGATTCCTTTGAA 640
Db      481  TCTTACTGGTATTTGCGGCGTCTCACCGAGTGGGATGCGAGATGATTCCTTTGAA 540
Qy      641  ATTAGGCGCTGAGATTCGCGCTTGTCTCTTGTCAAGCTGAATTCGATTCCTGC 700
Db      541  ATTAGGCGCTGAGATTCGCGCTTGTCTCTTGTCAAGCTGAATTCGATTCCTGC 600
Qy      701  CGAACGAAAAGAGATTCCTTCTTCTGCTGCTGCGAGTTTGAAGTTCATCTGCTT 760
Db      601  CGAACGAAAAGAGATTCCTTCTTCTGCTGCTGCGAGTTTGAAGTTCATCTGCTT 660
Qy      761  GTGATAATTCAGGTCAGGCTTATTTGCGGCTGCTGATCTTCTTGTGCTGTGACC 820
Db      661  GTGATAATTCAGGTCAGGCTTATTTGCGGCTGCTGATCTTCTTGTGCTGTGACC 720
Qy      821  ATCCGTAATCTTCTTGTGGAAGGCTGCTAAA 853
Db      721  ATCCGTAATCTTCTTGTGGAAGGCTGCTAAA 753

```

```

RESULT 4
US-09-471-803A-4
; Sequence 4, Application US/09471803A

```

```

; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PEEPERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLES OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708-8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(324)
; OTHER INFORMATION: brim
; OTHER INFORMATION: ATCC14752
US-09-471-803A-4

```

```

Query Match      25.5%; Score 324; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 8.9e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 853 ATGACACTGATTTCTCTGATTCCTGTTGTTGCGAGATGCGATCACTTCTTT 912
DB 1 ATGACACTGATTTCTCTGATTCCTGTTGTTGCGAGATGCGATCACTTCTTT 60
QY 913 GGGCTCCGGCGCGCTTCCTTTTAACTCTTAAGCCCCCTTACGATGCACTTTGGGC 972
DB 61 GGGCTCCGGCGCGCTTCCTTTTAACTCTTAAGCCCCCTTACGATGCACTTTGGGC 120
QY 972 AAAATGGCGATGATGATGCGACAGAGATCTTGGCATTTTACCGCATCACTTTGGC 1032
DB 121 AAAATGGCGATGATGATGCGACAGAGATCTTGGCATTTTACCGCATCACTTTGGC 180
QY 1032 AGCAATGCGATGATGATGAAAGCTTAACTTTGTTGTTGCGATTTACAGTG 1092
DB 181 AGCAATGCGATGATGATGAAAGCTTAACTTTGTTGTTGCGATTTACAGTG 240
QY 1093 GTGGCGCATCTTCTTGGCGGTGAGAGACACTTGTGAGGCTTGGCGCTGGCACATCGTT 1152
DB 241 GTGGCGCATCTTCTTGGCGGTGAGAGACACTTGTGAGGCTTGGCGCTGGCACATCGTT 300
QY 1153 TTTGTTGACTGATGATCTTTTC 1176
DB 301 TTTGTTGACTGATGATCTTTTC 324

```

```

RESULT 5
US-09-489-039A-3016
; Sequence 3016, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLES OF INVENTION: PNEUMONITIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3016
; LENGTH: 756
; TYPE: DNA

```

```

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

```

```

Query Match      4.6%; Score 58.6; DB 4; Length 756;
Best Local Similarity 52.2%; Pred. No. 3.4e-09;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

```

```

QY 195 TCGGGAAGGTCAAAAACCTCCCTGTCGACGTTTGGGCAATGACCGATGTATG 254
DB 62 TCGGGAAGGTCAAAAACCTCCCTGTCGACGTTTGGGCAATGACCGATGTATG 121
QY 255 CGTTGATCTCTGTTTATTAATGAGGCTAGCAATGATGGGCAAGCCCATGTTTCG 314
DB 122 CTTGCGACATCAATGCTACCCGCGCTGCTTCAACCCCGTGAAGCCGTTCTTCTCT 181
QY 315 GCCTGATTTTGGGGGCTCCACCGAATGCTGTATGCGCTTCGTTGGGGGCGAGGC 374
DB 182 GCATTAATTAAGCGCGCGCGACAGATTTGTATCACCCCATGTTGGCGCGGCGAGTT 241
QY 375 CCTGGGCGCATGCGGCTCACCAATGCTGTGTAATTCGCGCAAGTATTCTATGCGT 434
DB 242 CGCTGCGGTGCGCGCGCTGACGCGTATGCGGATGATGCGGCAAGTATGCGG 301
QY 435 TTTCATTC 443
DB 302 CTTCACTGC 310

```

```

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

IMMEDIATE SOURCE:
CLONE: PT29P-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.6%; Pred. No. 2e-06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

4.1%; Score 52.4; DB 1; Length 7218;
Query Match
Best Local Similarity 3.6%; Pred. No. 2e-06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

560 TCAATGCAATAGAGTTTCACTCTCTGAGTATCGGCGGCTCACCAGAGTGGCATC 619
1050 TCGAGGAGGCTTGGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1109
620 GCAAGATGATTCCTTTGAATTAAGGAGGCTCGAGGCTTGGCTCTCTTGTGC 679
1110 TT 1169
680 AGGCTGACTTGGATTCCTGCGAAGAGAGAGATCCCTCTGCTGCTGCAAGT 739
1170 TT 1229
740 TTGAGCTTCAACATTCCTCTGTGTATTCAGGTGAGGCTTGGCTCTGCTGCTG 799
1230 TT 1289
800 ATCTCTGAGGCTGTTGACCATCGGACTTCTTCTGGAAGGCTCTAATGACAA 859
1290 TT 1349
860 CTGATTTCTCTGATTCCTCTGTTGTCGAGATGTCAGTCAATTTTGGCTCC 919
1350 TT 1409
920 GGGCGGTTCCGTTCTTAATCCTTAAGCCCTA 951
1410 TTTTTTTTTTTTTTTTTTTTTTTTTTTTATCAAA 1441

RESULT 7
US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENE THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-09-107-532A-1266

Query Match
Best Local Similarity 3.1%; Score 39.8; DB 4; Length 768;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

207 TAAACCTCCCTGCTGAGGTTTGGGATGATCCCATTTGATTTGGCTCT 266
92 TCAAGACAGTTACTTACCTGTTTGGTTATATGATTTGACTTGTATGATG 151
267 TGGTATTCATATCGGCTACGAATGCTGGAGCCCATGTTTCCGGCTGATTTG 326
152 TTGGAAAGCTGCGGATTTTACATTCACATGCTGACGTTAGTCCCTGCTGCTATG 211
327 CGGCTCCACCGAAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
212 CTGATTCGCTGCAATTTATCAGTACGATGCTGCTGCTGCTGCTGCTGCTGCT 271
387 TCGGCTCACCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
272 TCGTTTCTGACATTTCTTACTGATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 302

RESULT 8
US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
NUMBER OF SEQUENCES: 107196.136
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11376
LENGTH: 1038
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11376

Query Match
Best Local Similarity 3.1%; Score 39.2; DB 4; Length 1038;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

309 TTTCCGCTGATTTTTCGCGGCTCCACGAATGCTGCTGCTGCTGCTGCTGCTG 368
473 TGTCCGCAATGCTTTCGCGGCGCGCCGCAATGCTGCTGCTGCTGCTGCTGCT 532
369 CAGGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 428
533 GTGCGGCTTCTTTGATCATCTTACACACCTGCTGCTGCTGCTGCTGCTGCT 592
429 ATGCGTTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 448

Db 593 ACGGATGAGCTTCGGGCG 612

RESULT 9

US-09-252-991A-11418
Sequence 11418, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11418

LENGTH: 1086

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11418

Query Match

Best Local Similarity 3.1%; Score 39.2; DB 4; Length 1086;
Pred. No. 0.015; Mismatches 63; Indels 0; Gaps 0;

Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

309 TTTCGGGCTGATTTTCGGGCTCCACGGAATGCTGATCGCCCTCGTTGGGCG 368

222 TTTCCGCTCCTGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 281

369 CAGCGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 428

282 GTGCGCGGCTTTCTTTGATCATCTTCACACCCCTCTCTGACCTGACGACCTGCT 341

429 ATGCGTTTCATTTCCCGCG 448

342 ACGGATGAGCTTCGGGCG 361

Db

RESULT 10

US-09-621-976-8976/C

Sequence 8976, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET 054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 8976

LENGTH: 399

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-8976

Query Match

Best Local Similarity 3.1%; Score 39; DB 4; Length 399;
Pred. No. 0.0086; Mismatches 97; Indels 0; Gaps 0;

Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

609 GAGTGGGATGCGAGATGATTCCTTTGAATTAAGGCGCTCGAGTCCCTTGGCT 668

242 SAARKKAKMTGKSRSSWSRSTGYAMTKSKSCTSRKMYKKRRKKRCKTSTXT 163

669 CTCTCTTGTACCGCTGACTTGTGATTCCTGCGGACGAAAGACAGATCCCTTCTGCG 728

Db

Query Match

Best Local Similarity 3.1%; Score 39; DB 4; Length 399;
Pred. No. 0.0086; Mismatches 97; Indels 0; Gaps 0;

Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

609 GAGTGGGATGCGAGATGATTCCTTTGAATTAAGGCGCTCGAGTCCCTTGGCT 668

242 SAARKKAKMTGKSRSSWSRSTGYAMTKSKSCTSRKMYKKRRKKRCKTSTXT 163

669 CTCTCTTGTACCGCTGACTTGTGATTCCTGCGGACGAAAGACAGATCCCTTCTGCG 728

Db

Query Match

Best Local Similarity 3.1%; Score 39; DB 4; Length 399;
Pred. No. 0.0086; Mismatches 97; Indels 0; Gaps 0;

Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

609 GAGTGGGATGCGAGATGATTCCTTTGAATTAAGGCGCTCGAGTCCCTTGGCT 668

242 SAARKKAKMTGKSRSSWSRSTGYAMTKSKSCTSRKMYKKRRKKRCKTSTXT 163

669 CTCTCTTGTACCGCTGACTTGTGATTCCTGCGGACGAAAGACAGATCCCTTCTGCG 728

Db 182 CTGAGTTCGKAAATYTKRRKRTWTYYYYSKMSKTKWKAATYATTKMWTETKWT 123

729 TGCTGACAGTTTGTAGCTTACCATGCTCTTGATTAATTCAGATGAGCCCTATTGG 788

122 CTMCWKCTTAAAGTMMYRARRYYAKAKMSKCTSTTCTYCKRYAKKCMSTYMSNS 63

789 CCGCGCTGCTGATCTTCTTGGGCTGTGACCATCGGACTTCTTGGGAAGGCT 847

62 MMKMSMMWMTYIYYTMMKSKTYYMSMSTYAKRKRYATYATYMTCTMTGKT 4

Db

RESULT 11

US-09-621-976-18033

Sequence 18033, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET 054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 18033

LENGTH: 474

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 16

OTHER INFORMATION: n=a, g, c or t

US-09-621-976-18033

Query Match

Best Local Similarity 2.9%; Score 37; DB 4; Length 474;
Pred. No. 0.045; Mismatches 156; Indels 0; Gaps 0;

Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

522 CGGCGAGCGCGAGGCTGCTGCGCGGCGGACTTATCTCAATGCAATAGCGTTCACT 561

13 YSGNCTCYCMARKMSYKXGSMYKSSCYKSCMCMCMKXCGSYKTTTWTAW 72

582 CTAAGTGTATTCGCGGCTCTCAGCGAGTGGCGAGTGGAGTGGATTCCTTTGAA 641

73 WYTTTGGKARBRMSGGKTYTMMCSKRTKSCMGRWKGYYSWYTCYKACTYM 132

642 TTAAGGCTCGAGTTCGCGCTTGTCTCTCTTGTCAAGTGAATTCCTGCG 701

133 WRRYVSSCCMNTYKGGGSMWTTTMMRRRKKSYKRTKGGKKTMMMAACYTWR 192

702 GAACGAAAAGACAGATCCCTTCTCTGCTGCTGCGAGTGGAGTGGATTCACATTCCTCTG 761

193 YMMMMRRRAAAKTYVCMMSKTMCMACCCMCMCMRRARSCCMRSRSTYMMCYTY 252

762 TGGTAATTCAGGTCAGGCGCTATTCGCGGCGTGGTGAATTCCTTGGCTGTGACCA 821

253 MMTGGRMYTMMRGGMMKMTMYTKKSMKSGCMRBAWARTTYTAMYYTYKR 312

822 TCCGTAATTCCTTCTTGGGAAAAGCTGCTAATGACACT 861

343 MCTYMRKTYCMWMMYSRWRSGMTRARGMWMMCYMY 352

Db

RESULT 12

US-09-252-991A-3205/C

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

US-09-252-991A-3205/C

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

US-09-252-991A-3205/C

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

US-09-252-991A-3205/C

Sequence 3205, Application US/09252991A

Patent No. 6551795

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3205
LENGTH: 1140
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3205

Query Match 2.9%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.097;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGACCTGATTTCGCGGCTCCACGAAATGCTGTCATCGCCCTGTTGGGCGCA 370
DB 465 TACTTCTGATCAACCTCGCGCGCCCTCGGCGCTGATCGGCTGACTCTCGGCGTC 406
QY 371 GCGCCCTGCGGCGCATCGGCTCACCACATTCGTTGTAAGTCCGCGACG 422
DB 405 GCGCCCTGCGGCGCATCTTCTGATCAGCGGCTGATCTTCTGTAAG 354

RESULT 13
US-09-252-991A-2761
Sequence 2761, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 2.9%; Score 36.8; DB 4; Length 2061;
Best Local Similarity 58.0%; Pred. No. 0.15;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGACCTGATTTCGCGGCTCCACGAAATGCTGTCATCGCCCTGTTGGGCGCA 370
DB 1273 TACTTCTGATCAACCTCGCGCGCCCTCGGCGCTGATCGGCTGACTCTCGGCGTC 1332
QY 371 GCGCCCTGCGGCGCATCGGCTCACCACATTCGTTGTAAGTCCGCGACG 422
DB 1333 GCGCCCTGCGGCGCATCTTCTGATCAGCGGCTGATCTTCTGTAAG 1384

RESULT 14
US-09-020-956-12/c
Sequence 12, Application US/09020956
Patent No. 6261562
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESSES:
ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-020-956-12

Query Match 2.9%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.099;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 327 CGGCTCCACGAAATGCTGTCATCGCCCTGTTGGGCGGCGGCGCGCA 386
DB 410 CGTTCTTCTTCAATCCTGATCTTCTGATGAGGTCAGTCTGTTGCGCT 351
QY 387 TCGCGCTCACCACATTCGTTGTAAGTCCGCGACGT 423
DB 350 TGGTGTACACACAAATGCTGAGACATCTTCTGAGCT 314

RESULT 15
US-09-030-607-12/c
Sequence 12, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESSES:
ADDRESSES: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-030-607-12

Query Match
Best Local Similarity 2.9%; Score 36.4; DB 3; Length 751;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 327 CCGGGCTCCACGGAATGCTGATCGCCCTGTTGTGGGCGCAGCGCCCTGGGGCCCA 386
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351

QY 387 TCGCGCTCACCACATGCTGTTGAACTTCGCGCAGCT 423
DB 350 TGGTGTACACACCAATGGCTGAGCACTTCCTGACGT 314

Search completed: April 19, 2004, 18:38:03
Job time : 113.018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 617.884 Seconds
(without alignments)
9204.305 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271

Sequence: 1 gcggcgatcaatgagatctctg.....agccttaacagacagcgct 1271

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCFUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	US-10-608-504-1	Sequence 1, Appl1
2	1271	100.0	1271	US-10-608-504-6	Sequence 6, Appl1
3	1271	100.0	3309400	US-09-738-626-1	Sequence 1, Appl1
4	753	59.2	753	US-09-738-626-289	Sequence 289, App
5	753	59.2	753	US-10-608-504-2	Sequence 2, Appl1
6	365.6	28.8	397	US-10-134-640-7	Sequence 7, Appl1
7	365.6	28.8	778	US-10-134-640-9	Sequence 9, Appl1
8	324	25.5	324	US-09-738-626-250	Sequence 250, App
9	324	25.5	324	US-10-608-504-4	Sequence 4, Appl1
10	185	14.6	715	US-10-134-640-1	Sequence 1, Appl1
11	127.4	10.0	711	US-09-738-626-3408	Sequence 3408, Ap
12	127.4	10.0	3309400	US-09-738-626-1	Sequence 1, Appl1
13	43	3.4	594	US-10-156-761-6067	Sequence 6067, Ap
14	43	3.4	9025608	US-10-156-761-1	Sequence 1, Appl1
15	41.4	3.3	351	US-10-184-644-80	Sequence 80, Appl1

C 16	41.4	3.3	351	14	US-10-184-634-80	Sequence 80, Appl1
C 17	40	3.1	671	14	US-10-184-644-346	Sequence 346, App
C 18	40	3.1	671	14	US-10-184-634-346	Sequence 346, App
C 19	39.8	3.1	596	14	US-10-184-644-310	Sequence 310, App
C 20	39.8	3.1	596	14	US-10-184-634-310	Sequence 310, App
C 21	39.8	3.1	596	14	US-10-063-665-100	Sequence 100, App
C 22	38.4	3.0	594	12	US-10-142-426-10	Sequence 10, Appl1
C 23	38.4	3.0	594	14	US-10-123-153-10	Sequence 10, Appl1
C 24	38.4	3.0	594	14	US-10-146-731-10	Sequence 10, Appl1
C 25	38.4	3.0	594	14	US-10-140-472-10	Sequence 10, Appl1
C 26	38.4	3.0	594	14	US-10-141-761-10	Sequence 10, Appl1
C 27	38.4	3.0	594	14	US-10-142-885-10	Sequence 10, Appl1
C 28	38.4	3.0	594	14	US-10-158-790-10	Sequence 10, Appl1
C 29	38.4	3.0	594	15	US-10-137-871-10	Sequence 10, Appl1
C 30	38.4	3.0	594	15	US-10-140-923-10	Sequence 10, Appl1
C 31	38.4	3.0	594	15	US-10-141-756-10	Sequence 10, Appl1
C 32	38.4	3.0	594	15	US-10-141-759-10	Sequence 10, Appl1
C 33	38.4	3.0	594	15	US-10-140-805-10	Sequence 10, Appl1
C 34	38.4	3.0	594	15	US-10-140-864-10	Sequence 10, Appl1
C 35	38.2	3.0	946	9	US-09-778-320-211	Sequence 211, App
C 36	38.2	3.0	946	9	US-09-910-689-211	Sequence 211, App
C 37	38.2	3.0	946	13	US-10-010-742-211	Sequence 211, App
C 38	37.8	3.0	457	14	US-10-184-644-48	Sequence 48, Appl1
C 39	37.8	3.0	457	14	US-10-184-634-48	Sequence 48, Appl1
C 40	37.8	3.0	457	14	US-10-063-665-12	Sequence 12, Appl1
C 41	37.4	2.9	557	14	US-10-184-644-326	Sequence 326, App
C 42	37.4	2.9	557	14	US-10-184-634-326	Sequence 326, App
C 43	37	2.9	186	14	US-10-156-761-5719	Sequence 5719, App
C 44	37	2.9	9025608	14	US-10-156-761-1	Sequence 1, Appl1
C 45	36.8	2.9	1230	12	US-10-282-122A-30075	Sequence 30075, A

ALIGNMENTS

RESULT 1
US-10-608-504-1
Sequence 1, Application US/10608504
Publication No. US20040014123A1
GENERAL INFORMATION:
APPLICANT: KENNERNECHT, NICOLE
APPLICANT: SAHM, HERMAN
APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 7601/80525
CURRENT APPLICATION NUMBER: US/10/608,504
CURRENT FILING DATE: 2003-06-30
PRIOR APPLICATION NUMBER: US 09/471,803
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnf
FEATURE:
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnf
FEATURE:
OTHER INFORMATION: ATCC14752
US-10-608-504-1
Query Match 100.0%, Score 1271, DB 15, Length 1271;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGGATCATGGAATCTGATTCATATATGSCAAATAGCCTAGTGAAGTGGCGCAAC 60
DB 1 GCGGATCATGGAATCTGATTCATATATGSCAAATAGCCTAGTGAAGTGGCGCAAC 60
QY 61 TGGCAACAAACTACCCGGCAATTGNGATGATGTGTGTGCAAAAACGCAAGAT 120
DB 61 TGGCAACAAACTACCCGGCAATTGNGATGATGTGTGTGCAAAAACGCAAGAT 120
QY 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACGATGATAAGTTA 180
DB 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACGATGATAAGTTA 180
QY 181 TGGGCGCTACGAATATGCGCGCAAGGTCTAAAACCTCCCTGTGTGAGGTTGGGATGTA 240
DB 181 TGGGCGCTACGAATATGCGCGCAAGGTCTAAAACCTCCCTGTGTGAGGTTGGGATGTA 240
QY 241 CCCGATTGGATTGCGTTGATCTCTTGTGTTATTCATATGCGCTACGAATGATGGCGCAG 300
DB 241 CCCGATTGGATTGCGTTGATCTCTTGTGTTATTCATATGCGCTACGAATGATGGCGCAG 300
QY 301 CCCGCTGTTTTCGCGCTGATTTTGGCGGGCTCCACCGAATGCTGTGATGCGCTTGT 360
DB 301 CCCGCTGTTTTCGCGCTGATTTTGGCGGGCTCCACCGAATGCTGTGATGCGCTTGT 360
QY 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGTAATTCGCGCA 420
DB 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGTAATTCGCGCA 420
QY 421 CGTATTCTATGCGTTTTCATTCGCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
DB 421 CGTATTCTATGCGTTTTCATTCGCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
QY 481 TTCGCTTTCGCGCTATGAGAGGCAACGCAATGCTGTGCGGCGAGCGCGCAGCTG 540
DB 481 TTCGCTTTCGCGCTATGAGAGGCAACGCAATGCTGTGCGGCGAGCGCGCAGCTG 540
QY 541 GTCGCGGTGCGCATCTATCTCAATGCAATAGCGTTTCACTCTGATGTTGCGGCG 600
DB 541 GTCGCGGTGCGCATCTATCTCAATGCAATAGCGTTTCACTCTGATGTTGCGGCG 600
QY 601 TCTCACCGGAGTGGCGCATGCGCAAGTGTATCTCTTTGAAATTAAGGCGCTGAGTGGC 660
DB 601 TCTCACCGGAGTGGCGCATGCGCAAGTGTATCTCTTTGAAATTAAGGCGCTGAGTGGC 660
QY 661 CCTTTGCTCTCTCTTTGTACGCTGACTTGTGATTCCTGCGCAAGAAACAGATCCC 720
DB 661 CCTTTGCTCTCTCTTTGTACGCTGACTTGTGATTCCTGCGCAAGAAACAGATCCC 720
QY 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTTGTGTAATTCAGGTCAGGC 780
DB 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTTGTGTAATTCAGGTCAGGC 780
QY 781 CCTATTGCGGCGCTGCTGATCTCTTGGGCTGTGAGCATCCGCTACTTCTCTTGGG 840
DB 781 CCTATTGCGGCGCTGCTGATCTCTTGGGCTGTGAGCATCCGCTACTTCTCTTGGG 840
QY 841 AAAAGCTGCTAAATGACATGATTTCTCTGTATTTCTCTTGTGTCGAGATGTGCA 900
DB 841 AAAAGCTGCTAAATGACATGATTTCTCTGTATTTCTCTTGTGTCGAGATGTGCA 900
QY 901 GTCAATTACTTTTGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGTGAATCA 960
DB 901 GTCAATTACTTTTGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGTGAATCA 960
QY 961 CAATTTGTGGGCAAAATGGCGATGTGATGCGCAGAGAAATCTTCCATTTTGACCGCA 1020
DB 961 CAATTTGTGGGCAAAATGGCGATGTGATGCGCAGAGAAATCTTCCATTTTGACCGCA 1020
QY 1021 TCAAGCTTTGGGCAAGATGCGATGATGTAATCTGAACCTCTTGTGTCATGCGCTT 1080
DB 1021 TCAAGCTTTGGGCAAGATGCGATGATGTAATCTGAACCTCTTGTGTCATGCGCTT 1080

```

```

DB 1021 TCAAGCTTTGGGCAAGATGCGATGATGTAATCTGAACCTCTTGTGTCATGCGCTT 1080
QY 1081 GCGATTACAGTGTGGCGCATCTTTGGGCGGTGACGCACTTGTGAGCGTTGGCGCT 1140
DB 1081 GCGATTACAGTGTGGCGCATCTTTGGGCGGTGACGCACTTGTGAGCGTTGGCGCT 1140
QY 1141 GGCACCATGCTTTTGTGTGATGTAATCTTTCTTAAACCTGATTAATAAATAAAT 1200
DB 1141 GGCACCATGCTTTTGTGTGATGTAATCTTTCTTAAACCTGATTAATAAATAAAT 1200
QY 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTTGAAGGCTTAAG 1260
DB 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTTGAAGGCTTAAG 1260
QY 1261 CAGACAGCGCT 1271
DB 1261 CAGACAGCGCT 1271

```

RESULT 2

```

US-10-608-504-6
; Sequence 6, Application US/10608504
; Publication No. US20040014123A1
; GENERAL INFORMATION:
; APPLICANT: KENNERNBECHT, NICOLE
; APPLICANT: SAMM, HERMANN
; APPLICANT: EGGELE, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; CURRENT FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; FEATURE:
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; FEATURE:
; OTHER INFORMATION: ATCC13032
US-10-608-504-6

```

Query Match 100.0%; Score 1271; DB 15; Length 1271;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGGATCATGGAATCTGATTCATATATGSCAAATAGCCTAGTGAAGTGGCGCAAC 60
DB 1 GCGGATCATGGAATCTGATTCATATATGSCAAATAGCCTAGTGAAGTGGCGCAAC 60
QY 61 TGGCAACAAACTACCCGGCAATTGNGATGATGTGTGTGCAAAAACGCAAGAT 120
DB 61 TGGCAACAAACTACCCGGCAATTGNGATGATGTGTGTGCAAAAACGCAAGAT 120
QY 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACGATGATAAGTTA 180
DB 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACGATGATAAGTTA 180
QY 181 TGGGCGCTACGAATATGCGCGCAAGGTCTAAAACCTCCCTGTGTGAGGTTGGGATGTA 240
DB 181 TGGGCGCTACGAATATGCGCGCAAGGTCTAAAACCTCCCTGTGTGAGGTTGGGATGTA 240

```

Db 181 TCGGCGCTACGAAATCGGCGAAGCTTAAACCTCCCTTCTCAGGTTGGGCAATGTA 240
Qy 241 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 300
Db 241 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 300
Qy 301 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 360
Db 301 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 360
Qy 361 TGTGGGCGAGGCGCCCTGGGCGCATGCGGCTCAACATTTGCTGTAATTCGCGCA 420
Db 361 TGTGGGCGAGGCGCCCTGGGCGCATGCGGCTCAACATTTGCTGTAATTCGCGCA 420
Qy 421 CGTATTCATGCGTTTCAATTCGCGCTCATGCTGTCATAAAACCCATTTGCGCTTCTA 480
Db 421 CGTATTCATGCGTTTCAATTCGCGCTCATGCTGTCATAAAACCCATTTGCGCTTCTA 480
Qy 481 TTGCGTTTTCGCGCTTATCGAGAAAGCTTACGATCTCGGCGCAGGCGCGAGGCTG 540
Db 481 TTGCGTTTTCGCGCTTATCGAGAAAGCTTACGATCTCGGCGCAGGCGCGAGGCTG 540
Qy 541 GTGCGCGTGGGAGCTTATCTCAATGCAATGAGCTTCTACTGCTGATTTGCGCGG 600
Db 541 GTGCGCGTGGGAGCTTATCTCAATGCAATGAGCTTCTACTGCTGATTTGCGCGG 600
Qy 601 TCTCACCGGAGTGGGAGTGGGAGTGGTATTTGAAATTAAGGCGCTGAGTTGCG 660
Db 601 TCTCACCGGAGTGGGAGTGGGAGTGGTATTTGAAATTAAGGCGCTGAGTTGCG 660
Qy 661 CCTTTGCTCTCTCTTGTCAAGCTGATTTGAAATTCCTGCGGAAAGGAAAGAGATCCC 720
Db 661 CCTTTGCTCTCTCTTGTCAAGCTGATTTGAAATTCCTGCGGAAAGGAAAGAGATCCC 720
Qy 721 TTCTCTGCTCTCGGAGGTTTGAAGCTTCAACATTCCTTGTGTAATTCAGGTTGAGG 780
Db 721 TTCTCTGCTCTCGGAGGTTTGAAGCTTCAACATTCCTTGTGTAATTCAGGTTGAGG 780
Qy 781 CCTATTGCGGCGCTGATCTTCTTGGGCTGTTGACCAACCGGATCTTCTTGGG 840
Db 781 CCTATTGCGGCGCTGATCTTCTTGGGCTGTTGACCAACCGGATCTTCTTGGG 840
Qy 841 AAAGGCTCTAAATGACAACTGATTTCTCTGATATTCCTTGTGTTGCGAGTATGCA 900
Db 841 AAAGGCTCTAAATGACAACTGATTTCTCTGATATTCCTTGTGTTGCGAGTATGCA 900
Qy 901 GTCATTTACTTTTGGCGCTCGGCGGCTTCCGTTCTTAATCCCTTAAGCCCTACGTAATCA 960
Db 901 GTCATTTACTTTTGGCGCTCGGCGGCTTCCGTTCTTAATCCCTTAAGCCCTACGTAATCA 960
Qy 961 CAATTTGTTGGGCAAAATGCGATGTGATGTCAGAGAAATCTTTCATTTTGAACGCA 1020
Db 961 CAATTTGTTGGGCAAAATGCGATGTGATGTCAGAGAAATCTTTCATTTTGAACGCA 1020
Qy 1021 TCAAGTTTTCGAGCAATGCGATGATGTAAGATCTTAACCTTTGCTGATTCGCGGT 1080
Db 1021 TCAAGTTTTCGAGCAATGCGATGATGTAAGATCTTAACCTTTGCTGATTCGCGGT 1080
Qy 1081 GCGATTACAGTGTGCGCATCTTCTTGGCGGTCAGCACTTGTGAAGGCTTGGCGCT 1140
Db 1081 GCGATTACAGTGTGCGCATCTTCTTGGCGGTCAGCACTTGTGAAGGCTTGGCGCT 1140
Qy 1141 GCGACCATCGTTTGTGAGTGTGATCTTTTCTAAACCTGATTAATAAAGCTTAAG 1200
Db 1141 GCGACCATCGTTTGTGAGTGTGATCTTTTCTAAACCTGATTAATAAAGCTTAAG 1200
Qy 1201 CCGGATGCGCTCAATTTGAAGGCGGATTTTGAAGGAACTTAAGAAAGGCTTAAG 1260
Db 1201 CCGGATGCGCTCAATTTGAAGGCGGATTTTGAAGGAACTTAAGAAAGGCTTAAG 1260
Qy 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

Db 1261 CAGACAGCGCT 1271
RESULT 3
US-09-738-626-1
Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 100.0%; Score 1271; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGGAGTCAATGGAATCTAGCTTCAATATTTGACAAATAGCTTATGAGTGGCGCAAC 60
Db 276729 GCGGAGTCAATGGAATCTAGCTTCAATATTTGACAAATAGCTTATGAGTGGCGCAAC 276788
Qy 61 TGGCAACAAATACCGCGCAATTTGTGATGATTTAGTGTGCAAAAAACGCAAGAT 120
Db 276789 TGGCAACAAATACCGCGCAATTTGTGATGATTTAGTGTGCAAAAAACGCAAGAT 276848
Qy 121 TCAATCAAGCTTGAAGTGTGCGATCCCAAGGAGCGCTGGAACCGATGATTAAGTTA 180
Db 276849 TCAATCAAGCTTGAAGTGTGCGATCCCAAGGAGCGCTGGAACCGATGATTAAGTTA 276908
Qy 181 TCGGCGCTAGGAATGCGCAAGGCTTAATAAACCCTTGTGCGAGTTTGGGCACTGA 240
Db 276909 TCGGCGCTAGGAATGCGCAAGGCTTAATAAACCCTTGTGCGAGTTTGGGCACTGA 276968
Qy 241 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 300
Db 276969 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 277028
Qy 301 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 360
Db 277029 CCCGATTGATATGAGTGTGCTCTTGTATTAATACGATACGATATGTTGGGAGC 277088
Qy 361 TGTGGGCGAGCGCCCTGGGCGCATGCGCTCAACATTTGCTGTAATTCGCGCA 420
Db 277089 TGTGGGCGAGCGCCCTGGGCGCATGCGCTCAACATTTGCTGTAATTCGCGCA 277148
Qy 421 CGTATTCATGCGTTTCAATTCGCGCTCATGCTGTCATAAAACCCATTTGCGCTTCTA 480
Db 277149 CGTATTCATGCGTTTCAATTCGCGCTCATGCTGTCATAAAACCCATTTGCGCTTCTA 277208
Qy 481 TTGCGTTTTCGCGCTTATCGAGAAAGCTTACGATCTCGGCGCAGGCGCGAGGCTG 540
Db 481 TTGCGTTTTCGCGCTTATCGAGAAAGCTTACGATCTCGGCGCAGGCGCGAGGCTG 540

```

Db 277209 TTGGTTTGGGCTTTATCGAGAGAGCTTACGAGCTCACTCGGCGGAGCCCGAGGCTG 277268
QY 541 GTGCGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGATTTCCGCGG 600
Db 277269 GTGCGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGATTTCCGCGG 277328
QY 601 TCTCACCGGAGTGGGAGATCGAGAGTTTATCTTCTTGAATTTAAGGCGCTCGAGTTCCG 660
Db 277329 TCTCACCGGAGTGGGAGATCGAGAGTTTATCTTCTTGAATTTAAGGCGCTCGAGTTCCG 277388
QY 661 CTTTGGCTCTCTCTTGTCAAGCTGACTTGTGATCTCTGCGGAAAGAAAGAGATCCC 720
Db 277389 CTTTGGCTCTCTCTTGTCAAGCTGACTTGTGATCTCTGCGGAAAGAAAGAGATCCC 277448
QY 721 TTTCTGCTGCTCGAGAGTTTGAAGTTTCAACCTTGTCTTGTGATTTCCAGGTCAGGC 780
Db 277449 TTTCTGCTGCTCGAGAGTTTGAAGTTTCAACCTTGTCTTGTGATTTCCAGGTCAGGC 277508
QY 781 CCTATTTGGGCGGCTGCTGATCTTCTGAGTCTGTGACCAATCCGAGTACTTCTTGGG 840
Db 277509 CCTATTTGGGCGGCTGCTGATCTTCTGAGTCTGTGACCAATCCGAGTACTTCTTGGG 277568
QY 841 AAAGCTGTGAATGACAACTGATTTCTCTGATTTCTCTTGTGTCAGATGTCGA 900
Db 277569 AAAGCTGTGAATGACAACTGATTTCTCTGATTTCTCTTGTGTCAGATGTCGA 277628
QY 901 GTGATTTAGGCTCGGCTCGGCGGCTCGGCTCTTAATCTCTTAAGCCCTTAAGTGAATCA 960
Db 277629 GTGATTTAGGCTCGGCTCGGCGGCTCGGCTCTTAATCTCTTAAGCCCTTAAGTGAATCA 277688
QY 961 CAATTTGTGGGCAAAATGGCGATGTGATGCCAGAGGAATCTTGCATTTTGAACGCA 1020
Db 277689 CAATTTGTGGGCAAAATGGCGATGTGATGCCAGAGGAATCTTGCATTTTGAACGCA 277748
QY 1021 TTAAGCTTTCGAGCAATAGCATATCTGAAGATCTTAACCTTGTGTCATTTGCCCTT 1080
Db 277749 TTAAGCTTTCGAGCAATAGCATATCTGAAGATCTTAACCTTGTGTCATTTGCCCTT 277808
QY 1081 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACCTTGTGAGGCTTGGCGCT 1140
Db 277809 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACCTTGTGAGGCTTGGCGCT 277868
QY 1141 GCGACCATGCTTTTGTGAGTGTGAATCTTTTCTAATAATGCAATTAATCAAAAT 1200
Db 277869 GCGACCATGCTTTTGTGAGTGTGAATCTTTTCTAATAATGCAATTAATCAAAAT 277928
QY 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
Db 277929 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 277988
QY 1261 CAGACAGCGCT 1271
Db 277989 CAGACAGCGCT 277999

```

```

RESULT 4
US-09-738-626-289
; Sequence 289, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125

```

```

; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 289
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-289

```

```

Query Match 59.2%; Score 753; DB 9; Length 753;
Best Local Similarity 100.0%; Pred. No. 6,5e-245; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

```

```

QY 101 GTGCAAAAAAGCAAGAGATTCAATCAAGCTGAGAGTGTGCGCATCCAGGCAAGCCCTG 160
Db 1 GTGCAAAAAAGCAAGAGATTCAATCAAGCTGAGAGTGTGCGCATCCAGGCAAGCCCTG 60
QY 161 GAACGAGATGATTAAGGTTATGCGGCTGACGAATATGCGCAAGGTTAAACCTTCCCTT 220
Db 61 GAACGAGATGATTAAGGTTATGCGGCTGACGAATATGCGCAAGGTTAAACCTTCCCTT 120
QY 221 GCTGCAAGTTTGGGCAATGATACCGGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 280
Db 121 GCTGCAAGTTTGGGCAATGATACCGGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 180
QY 281 GAGTACGATGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 340
Db 181 GAGTACGATGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 341 ATGCTGATATGCGGCTGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
Db 241 ATGCTGATATGCGGCTGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 401 TTGCTGATGATTTGCGGCAAGTATCTATGCTTTTCAATCCGCTGATGATGATGATGATGAT 460
Db 301 TTGCTGATGATTTGCGGCAAGTATCTATGCTTTTCAATCCGCTGATGATGATGATGATGAT 360
QY 461 AAGCCCATTTGCGGCTTCTATTTGCTTTTGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAG 520
Db 361 AAGCCCATTTGCGGCTTCTATTTGCTTTTGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 521 GCGGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
Db 421 GCGGCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 581 TCTTACTGAGGATTTGCGGCTTCTACCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 640
Db 481 TCTTACTGAGGATTTGCGGCTTCTACCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
QY 641 ATTAAGGAGCTGAGGCTTCTGCTTCTCTTGTGACGCTGATTTGATTTGATTTGATTTGAT 700
Db 541 ATTAAGGAGCTGAGGCTTCTGCTTCTCTTGTGACGCTGATTTGATTTGATTTGATTTGAT 660
QY 701 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 760
Db 601 GAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 761 GTGATATTCAGGATGAGGCTTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 820
Db 661 GTGATATTCAGGATGAGGCTTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 821 ATCCGATCTTCTTCTTGGGAAAGGCTGCTAA 883
Db 721 ATCCGATCTTCTTCTTGGGAAAGGCTGCTAA 783

```



```

; APPLICANT: SAEW, HERMANN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 S1 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(324)
; OTHER INFORMATION: brne
; FEATURE:
; OTHER INFORMATION: ATCC14752
; US-10-608-504-4

Query Match      25.5%; Score 324; DB 15; Length 324;
Best Local Similarity 100.0%; Pred. No. 4,3e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      883 ATGACACTGATTTCTCTCTGATTTCTCTCTGTTGTCGAGATGTCAGTCAATTGCTTT 912
DB      1  ATGACAACTGATTTCTCTCTGATTTCTCTCTGTTGTCGAGATGTCAGTCAATTGCTTT 60

QY      913 GCGCTCCGGGCGGCTTCCGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 972
DB      61 GCGCTCCGGGCGGCTTCCGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 120

QY      973 AAAATGCGGATGATGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1032
DB      121 AAAATGCGGATGATGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 180

QY      1033 AGCAATGCGATGATGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 1092
DB      181 AGCAATGCGATGATGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 240

QY      1093 GTGGCGCATCTTCTTGGCGGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTT 1152
DB      241 GTGGCGCATCTTCTTGGCGGTCGAGACGATCTTAACTTAACTTAACTTAACTTAACTTAACTT 300

QY      1153 TTTGTTGACCTGTGTAATCTTTTC 1176
DB      301 TTTGTTGACCTGTGTAATCTTTTC 324

RESULT 10
US-10-134-640-1/c
; Sequence 1, Application US/10134640
; Publication No. US20030017555A1
; GENERAL INFORMATION:
; APPLICANT: Bathe, Brigitte
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Puhler, Alfred
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfefferle, Walter
; TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
; FILE REFERENCE: 990109 BT
; CURRENT APPLICATION NUMBER: US/10/134,640
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```

; LENGTH: 715
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(612)
; OTHER INFORMATION: lrp-Gen
; FEATURE:
; NAME/KEY: -10 signal
; LOCATION: (88)..(93)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: -35 signal
; LOCATION: (62)..(67)
; OTHER INFORMATION:
; US-10-134-640-1

Query Match      14.6%; Score 185; DB 12; Length 715;
Best Local Similarity 100.0%; Pred. No. 1,5e-51;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GCGGATCAATGGAATCTAGCTCATATATTCACAAATAGCCTAGTTGAGGTGCGAAAC 60
DB      185 GCGGATCAATGGAATCTAGCTCATATATTCACAAATAGCCTAGTTGAGGTGCGAAAC 126

QY      61  TGGCAACAAACTAACCCGCAATTTGTGTGATGATTTGTAGTGTGCAAAAAACGCAAGAT 120
DB      125 TGGCAACAAACTAACCCGCAATTTGTGTGATGATTTGTAGTGTGCAAAAAACGCAAGAT 66

QY      121 TCATTCAAGCTGAGAGGTGCGCATTCGCAAGGAGCCCTGGAACGATGATTAAGGTTA 180
DB      65 TCATTCAAGCTGAGAGGTGCGCATTCGCAAGGAGCCCTGGAACGATGATTAAGGTTA 6

QY      181 TCGGC 185
DB      5 TCGGC 1

RESULT 11
US-09-738-626-3408
; Sequence 3408, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENO, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3408
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-3408

Query Match      10.0%; Score 127.4; DB 9; Length 711;
Best Local Similarity 51.5%; Pred. No. 6,1e-32;
```

```

: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 1
: LENGTH: 3309400
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match          10.0%; Score 127.4; DB 9; Length 3309400;
Best Local Similarity 51.5%; Pred. No. 1,4e-29;
Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 222 CTGCAAGTTGGGCATGTACCCGATGTGATATGCGTTGTGCTCTCTTGATTCATATACG 281
DB 3289969 CGTCCGATTGGGGCTATTTCGCTGGATTGGCCATTGAGTCCACAG 3289910
QY 282 GCTACGATGGTTGGGACGCCCATCTGTTTCCGGCCATGTTTCGCGAGCTCCACCGAA 341
DB 3289909 GTTTCGCTGTGTGTGACGCCCATTTTCTCTTCGTGATCTATGCGGATTGATGAA 3289850
QY 342 TGCTGTCATCGCCCTGTGTGTGGGCGACGCGCCCTTGAGGCGCATGCGCTACCAAT 401
DB 3289849 TTCTGGCATCGCATGATCAACCCAGATACGCGCCGTTTTCGCGCGGCTGCTGAT 3289790
QY 402 TGCTGTGAATTCGCCGACAGTATTCATATGCGTTTCATTCGCCGTCATGTGTCAAA 461
DB 3289789 TCATGTGAATTCGCCGACATTTTCTACGCTTCACCTTCCACGCCACCGCATCAAG 3289730
QY 462 ACCCATTTGCCCGTTTCTATTGCGTTTTCGCGCTTATCGAAGAGCTTACGAGTCACTG 521
DB 3289729 CCGGCGCGCGCGCGGCTATTCCACTACGAGGCTTACCGAGATCTTACGCCATG 3289670
QY 522 CGGCGAGGCCCGCAGGCTGTGCGCGG---GGCACTTATCTCATAGCAATAGCGTTTC 578
DB 3289669 CAGCCCGGCCACCTGGCGCATATCACTGCGACGCGGAGCTTACCGTTCAATTTTGTGCC 3289610
QY 579 ACTCTCACTGGGATATTCGCGCGCTCTCAACCGAGTGGCGATCGCAGATTGATTCCTTTG 638
DB 3289609 AAGCTCTGTGGATTATCCACAGAAATATATGGCGGCTTGTGTGATGATGTCGCCGAT 3289550
QY 639 AAATTAAGGCGCTGAGTTCGCGCTTCTCTCTTGTGTACGCTGACTTGGATTCCT 698
DB 3289549 ATCTAAAGGCGATGATATTTCCCTGACCGGCGCTGTTGTGTGCTGCGGAGGAT 3289490
QY 699 GCGCGACAAAGAGCATCCCTCTCTGCTGCTGCGAGTTGAGCTTCAACATTGCTC 758
DB 3289489 TCATAAATTAACAAGATTATTCGCTGCCATTATTCGCGGTGATTTGGCTCTGATTCG 3289430
QY 759 TTGTGTAATTCAGGTGAGGCCCTTATTTGGCGGCTGCTGATCTTTGAGGCTGTGA 818
DB 3289429 GTTTGTGCGCGCGAGCAGATGCTGATTATCGCTTTGACACGATTTTGTGATCTTC 3289370
QY 819 CCATCCGCTACTTCTTC 835
DB 3289369 TTCTCCGCGTCCGCTTC 3289353

RESULT 13
US-10-156-761-6067
: Sequence 6067, Application US//10156761
: Publication No. US20030119018A1
: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, NASAHIRA
: TITLE OF INVENTION: NOVEL POLYPEPTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156,761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089

```

PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6067
LENGTH: 594
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-10-156-761-6067

Query Match 3.4%; Score 43; DB 14; Length 594;
Best Local Similarity 50.7%; Pred. No. 0.0029;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCGGGGCTCCAGCAAAATGCTGATCGCCCTCGTTGTGGGCGCAGCGCCCTGCGGC 383
DB 203 TCACCGGCTGGCCATCGCGCAGGTGCTCGGCATGTTGTGGCAGCGCGCTCGGCTGG 262
QY 384 CCATCGCGCTCAACCAATGCTGTGAACTTCCGCCAGTATTGTATGCGTTTCATTCC 443
DB 263 GCAACGTGCCAGCAGATCTCGGCGATGCTGCTGCTTCTTCTGCGGCTACTGCTCA 322
QY 444 CGCTGCATGTGTCAAAACCCCAATTGCGGCTTCTATGCTTTTGGCGCTTATGACG 503
DB 323 CGTTGCGCGAGTGTGTGAAGCGCGGTGTCGCGCTTCGCGACCGCCTTGCGGCTCG 382
QY 504 AAGCTACGCACTCACTGCGGCC 526
DB 383 CCGCGACACCCCTGTCATCGCC 405

RESULT 14
US-10-156-761-1/c

Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.4%; Score 43; DB 14; Length 9025608;
Best Local Similarity 50.7%; Pred. No. 1.4;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCGGGGCTCCAGCAAAATGCTGATCGCCCTCGTTGTGGGCGCAGCGCCCTGCGGC 383
DB 7327430 TCACCGGCTGGCCATCGCGCAGGTGCTCGGCATGTTGTGGCAGCGCGCTCGGCTGG 7327371

QY 384 CCATCGCGCTCAACCAATGCTGTGAACTTCCGCCAGTATTGTATGCGTTTCATTCC 443
DB 7327370 GCAACGTGCCAGCAGATCTCGGCGATGCTGCTGCTTCTTCTGCGGCTACTGCTCA 7327311
QY 444 CGCTGCATGTGTCAAAACCCCAATTGCGGCTTCTATGCTTTTGGCGCTTATGACG 503
DB 7327310 CGTTGCGCGAGTGTGTGAAGCGCGGTGTCGCGCTTCGCGACCGCCTTGCGGCTCG 7327251
QY 504 AAGCTACGCACTCACTGCGGCC 526
DB 7327250 CCGCGACACCCCTGTCATCGCC 7327228

RESULT 15
US-10-184-644-80/c

Sequence 80, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 80
LENGTH: 351
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-80

Query Match 3.3%; Score 41.4; DB 14; Length 351;
Best Local Similarity 13.4%; Pred. No. 0.0073;
Matches 45; Conservative 103; Mismatches 187; Indels 0; Gaps 0;

QY 161 GAACCAATGATTAAGGTTATCGCGCTCAAGAAATCGCGCAGGTCAAAACCTCCCTT 220
DB 350 GAD..B.Y.G.YGMB.GGMD.MSGV.T..B.TAD.CYCG..GCH.TMSANGAICYA 291
QY 221 GCTGCAAGTTGGGCAATGACCGAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
DB 290 C.BESYGH..G.H.S..RB.R.HAD.M..TNY.B..TYSSCBY..B.ATCH.K.M.TDC 231
QY 281 GGCTACGATGTGGGCGCCCACTGTTTCCGCGCTGATTTTCCGCGGCTCCACGGA 340
DB 230 B.Y..BTRYGMB.GSCBDCDMG.BYKDA..TWYCT.NND.NK.TYSSCTMSY.YBH 171
QY 341 ATGCTGTCATCGCCCTGCTGTGTTGGGCGCAGCGCCCTGCGCCATCGGCTCACCACA 400
DB 170 B.S.S.TBCRT.NHSGSW..C..SBDBAHVHCGM...CSGTVAAT.TBCTSS.TRB. 111
QY 401 TTGCTGTGTAATTCGCCAGCATGATTTGATGCGTTTTCATTTCCGCGCTGATGTGCAAA 460
DB 110 TT.YAC.ABMYC.B..SH.ASGMYNY..R.G..T..TCYBSHMB..NMGK.B.Y. 51
QY 461 AACCACATGCGCGTTCTATGCTGCTGCGCT 495
DB 50 ..CM.M.GA...S.SCBS.MT.R.MNSTITTS.BT 16

Search completed: April 17, 2004, 21:00:17
Job time: 647.884 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 4560.51 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176

Perfect score: 1076

Sequence: 1 gtgcaaaaacgaagcagat.....ctgactggtgactcttc 1076

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sbs:*
28: em_un:*
29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rtd:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_by:*
39: em_hgco_hum:*
40: em_hgco_mus:*
41: em_hgco_other:*

score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	100.0	1271	6	AR391953 Sequence
2	1076	100.0	1271	6	AR391956 Sequence
3	1076	100.0	1271	6	AX137709 Sequence
4	1076	100.0	1271	6	AX137714 Sequence
5	1076	100.0	1271	6	BD014991 Nucleotide
6	1076	100.0	1271	6	BD014994 Nucleotide
7	1076	100.0	2105	1	AF454053 Coryneb
8	1076	100.0	340000	1	AP005274 Coryneb
9	1076	100.0	349800	6	AX120085 Sequence
10	753	70.0	753	6	AR391954 Sequence
11	753	70.0	753	6	AX120373 Sequence
12	753	70.0	753	6	AX137710 Sequence
13	753	70.0	753	6	BD014992 Nucleotide
14	753	70.0	753	6	BD162490 Nucleotide
15	324	30.1	324	6	AR391955 Sequence
16	324	30.1	324	6	AX120374 Sequence
17	324	30.1	324	6	AX137712 Sequence
18	324	30.1	324	6	BD014993 Nucleotide
19	324	30.1	324	6	BD162491 Nucleotide
20	273.6	25.4	397	6	AX137083 Sequence
21	273.6	25.4	397	6	AX137085 Sequence
22	149.6	13.9	87340	1	AP005224 Coryneb
23	132.8	12.3	53793	1	BX248361 Coryneb
24	127.4	11.8	711	6	AX123492 Sequence
25	127.4	11.8	711	6	BD165609 Novel pol
26	127.4	11.8	309400	6	AX127153 Sequence
27	127.4	11.8	325651	1	AP005283 Coryneb
28	85	7.9	715	6	AX137077 Sequence
29	85	7.9	715	6	BD013995 Novel nuc
30	59.8	5.6	349926	1	BX571660 Wolinella
31	58.6	5.4	756	6	AR386287 Sequence
32	57.2	5.3	9888	1	AE011050 Methanosa
33	54.4	5.1	856	9	HSB25448 Homo sapi
34	52.4	4.9	7218	1	166494 Sequence 14
35	51.8	4.8	11025	1	AE013585 Methanosa
36	51.2	4.8	125020	9	AF429315 Homo sapi
37	51	4.7	308015	1	AE016783 Pseudomon
38	48	4.5	10732	6	E32986 Gene encodi
39	48	4.5	301838	1	AE017209 Geobacter
40	47	4.4	2000	6	AX655393 Sequence
41	46.6	4.3	3619	1	ECB410307 Erwinia c
42	46.6	4.3	10773	1	AE013695 Yersinia
43	46.6	4.3	220050	1	AJ414156 Yersinia
44	46	4.3	215050	1	AL646084 Ralstonia
45	45.8	4.3	20342	1	AE008888 Salmonella

ALIGNMENTS

RESULT 1
AR391953
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
AUTHORS 1 (bases 1 to 1271)
TITLE Kennertknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
Nucleotide sequences coding for the export of branched chain amino
acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;

FEATURES

source

Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 160
QY 61 GAAACGATGATTAAGATTATCGGCGCTACGAAATGCGGCAAGGTCTAAAACTCCCTT 120
DB 161 GAAACGATGATTAAGATTATCGGCGCTACGAAATGCGGCAAGGTCTAAAACTCCCTT 220
QY 121 GCTGAGGTTTGGGCAATGATACCGATTGGATTCGTTGATCTCTGTTATTCATATC 180
DB 221 GCTGAGGTTTGGGCAATGATACCGATTGGATTCGTTGATCTCTGTTATTCATATC 280
QY 181 GGTACGAAATGATGAGGCGAGCCCACTGTTTCCGCTGATATTTGCGGGCTCCACCGAA 240
DB 281 GGTACGAAATGATGAGGCGAGCCCACTGTTTCCGCTGATATTTGCGGGCTCCACCGAA 340
QY 241 ATGCTGTATATGCGGCTGCTGTTGAGGCGAGGCGCCCTGAGGCGCATCGGCTCACACA 300
DB 341 ATGCTGTATATGCGGCTGCTGTTGAGGCGAGGCGCCCTGAGGCGCATCGGCTCACACA 400
QY 301 TTGCTGTGAATCTTCGCAAGTATTCTATGCTTTTCAATCCGCTGATGTGTGAA 360
DB 401 TTGCTGTGAATCTTCGCAAGTATTCTATGCTTTTCAATCCGCTGATGTGTGAA 460
QY 361 AACCCATTGCGCGTTTCTATGAGTTTCGCGCTATGAGAGAACCTACAGAGTCACT 420
DB 461 AACCCATTGCGCGTTTCTATGAGTTTCGCGCTATGAGAGAACCTACAGAGTCACT 520
QY 421 GCGGCGAGGCGCGGAGGCTGTGCGGCTGCGCATTAATCATAGCAATAGCGTTTCA 480
DB 521 GCGGCGAGGCGCGGAGGCTGTGCGGCTGCGCATTAATCATAGCAATAGCGTTTCA 580
QY 481 TCTTACTGGGATTCGCGGCTTCAACGAGTGGCGATGCGCAAGTTGATTCCTTTGAA 540
DB 581 TCTTACTGGGATTCGCGGCTTCAACGAGTGGCGATGCGCAAGTTGATTCCTTTGAA 640
QY 541 ATTAAGGAGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTG 600
DB 641 ATTAAGGAGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTG 700
QY 601 CGAAGCAAAAACGATATCCCTTCTCTGCTGCTGCGAGTTGAGCTTACACATTCCTT 660
DB 701 CGAAGCAAAAACGATATCCCTTCTCTGCTGCTGCGAGTTGAGCTTACACATTCCTT 760
QY 661 GTGGTATTCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATCTCTTGGGCTGTGACC 720
DB 761 GTGGTATTCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATCTCTTGGGCTGTGACC 820
QY 721 ATCGGTAATCTTCTTTGGGAAAGCTGTCTAATGACATGATTTCTCTGATTTCTG 780
DB 821 ATCGGTAATCTTCTTTGGGAAAGCTGTCTAATGACATGATTTCTCTGATTTCTG 880
QY 781 TTGTTGTCGAGTATGTCAGGCTCATTTCTTTGCGCTCGGGGCGGTTCCGTTCTATATC 840
DB 881 TTGTTGTCGAGTATGTCAGGCTCATTTCTTTGCGCTCGGGGCGGTTCCGTTCTATATC 940
QY 841 TTAAAGCCCTTACGTAATCAATTTTGGGCAAAATGCGATGTGATCCGACAGGAA 900
DB 941 TTAAAGCCCTTACGTAATCAATTTTGGGCAAAATGCGATGTGATCCGACAGGAA 1000
QY 901 TCCCTGCAATTTTACCGGATCAAGCTTTGCGAGCAATGCGATGATCTGAAGCTCTAA 960
DB 1001 TCCCTGCAATTTTACCGGATCAAGCTTTGCGAGCAATGCGATGATCTGAAGCTCTAA 1060

```

```

QY 961 CCTTGTGATTCATTCGCGTTTGGCATTAACAGTGTGCGGATCTTCTGGCGGTCAGCGA 1020
DB 1061 CCTTGTGATTCATTCGCGTTTGGCATTAACAGTGTGCGGATCTTCTTGGCGGTCAGCGA 1120
QY 1021 CCTTGTGATTCATTCGCGTTTGGCATTAACAGTGTGCGGATCTTCTTGGCGGTCAGCGA 1076
DB 1121 CCTTGTGATTCATTCGCGTTTGGCATTAACAGTGTGCGGATCTTCTTGGCGGTCAGCGA 1176

```

RESULT 2

AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

Unknown.

Unclassified.

1 (bases 1 to 1271)

Kennerknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.

Nucleotide sequences coding for the export of branched chain amino

acids, process for the isolation thereof and use thereof

Patent: US 6613545-A 6 02-SEP-2003;

JOURNAL

FEATURES

Location/Qualifiers

source

ORIGIN

1..1271
/organism="unknown"
/mol_type="genomic DNA"

Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 160
QY 61 GAAACGATGATTAAGATTATCGGCGCTACGAAATGCGGCAAGGTCTAAAACTCCCTT 120
DB 161 GAAACGATGATTAAGATTATCGGCGCTACGAAATGCGGCAAGGTCTAAAACTCCCTT 220
QY 121 GCTGAGGTTTGGGCAATGATACCGATTGGATTCGTTGATCTCTGTTATTCATATC 180
DB 221 GCTGAGGTTTGGGCAATGATACCGATTGGATTCGTTGATCTCTGTTATTCATATC 280
QY 181 GGTACGAAATGATGAGGCGAGCCCACTGTTTCCGCTGATATTTGCGGGCTCCACCGAA 240
DB 281 GGTACGAAATGATGAGGCGAGCCCACTGTTTCCGCTGATATTTGCGGGCTCCACCGAA 340
QY 241 ATGCTGTATATGCGGCTGCTGTTGAGGCGAGGCGCCCTGAGGCGCATCGGCTCACACA 300
DB 341 ATGCTGTATATGCGGCTGCTGTTGAGGCGAGGCGCCCTGAGGCGCATCGGCTCACACA 400
QY 301 TTGCTGTGAATCTTCGCAAGTATTCTATGCTTTTCAATCCGCTGATTTGATTCCTG 360
DB 401 TTGCTGTGAATCTTCGCAAGTATTCTATGCTTTTCAATCCGCTGATTTGATTCCTG 460
QY 421 GCGGCGAGGCGCGGAGGCTGTGCGGCTGCGCATTAATCATAGCAATAGCGTTTCA 480
DB 521 GCGGCGAGGCGCGGAGGCTGTGCGGCTGCGCATTAATCATAGCAATAGCGTTTCA 580
QY 481 TCTTACTGGGATTCGCGGCTTCAACGAGTGGCGATGCGCAAGTTGATTCCTTTGAA 540
DB 581 TCTTACTGGGATTCGCGGCTTCAACGAGTGGCGATGCGCAAGTTGATTCCTTTGAA 640
QY 541 ATTAAGGAGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTG 600

```

Db 641 ATTAAGGGCCCTGAGTTGCGCCCTTGGCTCTCTTTGTGACAGCTGACTTTGATTCCTGC 700
 QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCCAGAGTTTGAAGCTTACCACTTCTCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCCAGAGTTTGAAGCTTACCACTTCTCTT 760
 QY 661 GTGGTAATTCAGAGTACAGGCGCCCTATTTGGCGCGCTGCGATCTCTTGGGCTGTGAC 720
 Db 761 GTGGTAATTCAGAGTACAGGCGCCCTATTTGGCGCGCTGCGATCTCTTGGGCTGTGAC 820
 QY 721 ATCCGTAATCTCTTCTTGGGAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 780
 Db 821 ATCCGTAATCTCTTCTTGGGAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 880
 QY 781 TTGTTGTCGAGTATGTCAGTATCTTTTGGGCTCCGGCGGCTCCGTTCTTATCC 840
 Db 881 TTGTTGTCGAGTATGTCAGTATCTTTTGGGCTCCGGCGGCTCCGTTCTTATCC 940
 QY 841 TTAAGCCCTACGTGAATCAATTTGTGGCAAAATGGCGATGTGATGCGACAGAGAA 900
 Db 941 TTAAGCCCTACGTGAATCAATTTGTGGCAAAATGGCGATGTGATGCGACAGAGAA 1000
 QY 901 TCCTTGGCATTTTGAACCGCATCAACGTTTGGCAGCAATGCGATGATCTGAAGACTTAA 960
 Db 1001 TCCTTGGCATTTTGAACCGCATCAACGTTTGGCAGCAATGCGATGATCTGAAGACTTAA 1060
 QY 961 CCTTGGTCTCATTCGCGCTGCGATTAACAGTGTGGCGCATCTTCTGGCGGTGACGCA 1020
 Db 1061 CCTTGGTCTCATTCGCGCTGCGATTAACAGTGTGGCGCATCTTCTGGCGGTGACGCA 1120
 QY 1021 CCTTGGTGGCGCTTGGCGCTGCGACCATGTTTTTGTGAGCTGTGAATCTTTTC 1076
 Db 1121 CCTTGGTGGCGCTTGGCGCTGCGACCATGTTTTTGTGAGCTGTGAATCTTTTC 1176

RESULT 3
 AX137709 1271 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137709
 DEFINITION Sequence 1 from Patent EP1096010.
 ACCESSION AX137709
 VERSION AX137709.1 GI:14273886
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 AUTHORS Kennerknecht, N., Eggeling, L., Sahm, H., and Pfeifferle, W.
 TITLE Nucleotide sequences coding for branched-chain amino acids export
 JOURNAL proteins, method for isolating them and their use
 Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 FEATURES
 source
 1. 1271
 Location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC4752"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="REMBL:CAC39923"
 /translation="MKTQELHSLSEVPSRALEPPDKGRYREINAGKTSIAGL
 GMYPIGIFGLVLIQYGEWMAAPLPSGLIFAGSTEMVLAVLVGAAPLGAIALTTLL
 NVFRVFAFSPFLVNVNPIARFYSVALIDEAVAVTAAPAPAGSAWRLISQIASH
 SYVVGGLGVIAELIPEIKGLEFALCSLFTVLTLDSCRTKKQIPSLIAGSFTI
 ALVVPQALPAAALIFGLLTIRYFGRKAK"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1

ORIGIN

Query Match 100.0%; Score 1076; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 8.3e-275;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAGAGATTCATTCAGAGCTGAGAGTGTGCGATCCAGAGCAGCCCTG 60
 Db 101 GTGCAAAAAGCAGAGATTCATTCAGAGCTGAGAGTGTGCGATCCAGAGCAGCCCTG 160
 QY 61 GAACGATGATPAAAGTTATCGCGCTACGAAATCGCGCAAGTCTPAAAACCTCCCTT 120
 Db 161 GAACGATGATPAAAGTTATCGCGCTACGAAATCGCGCAAGTCTPAAAACCTCCCTT 220
 QY 121 GGTGCAAGTTTGGGCAATGACCCGATGTGATTTGGGTTGCTCTGTGTTATTAATAC 180
 Db 221 GGTGCAAGTTTGGGCAATGACCCGATGTGATTTGGGTTGCTCTGTGTTATTAATAC 280
 QY 181 GGTGCAAGTTTGGGCAAGCCCACTGTTTTCGCGCTGATTTTGGCGGCTTCAACGAA 240
 Db 281 GGTGCAAGTTTGGGCAAGCCCACTGTTTTCGCGCTGATTTTGGCGGCTTCAACGAA 340
 QY 241 ATGCTGTATATGCTCTGTTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTCAACGA 300
 Db 341 ATGCTGTATATGCTCTGTTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTCAACGA 400
 QY 301 TTGCTGTGAATCTTCGCGACAGATTTCTATGCGTTTCAATCCCGTGCATGTGTCAA 360
 Db 401 TTGCTGTGAATCTTCGCGACAGATTTCTATGCGTTTCAATCCCGTGCATGTGTCAA 460
 QY 361 AACCCATTTGCCCTTTCTATTCGCTTTTGGGCTTATGAGAGAACCTTACGACGACT 420
 Db 461 AACCCATTTGCCCTTTCTATTCGCTTTTGGGCTTATGAGAGAACCTTACGACGACT 520
 QY 421 GCGGACAGGCGCGCAGAGTGTGCGGCTGCGACATTAATCTCATGCAATATAGCGTTTCA 480
 Db 521 GCGGACAGGCGCGCAGAGTGTGCGGCTGCGACATTAATCTCATGCAATATAGCGTTTCA 580
 QY 481 TCCTACTGAGTATTCGCGGCTTACCGGATGCGAGTGTGCAAGTGAATTCCTTTGAA 540
 Db 581 TCCTACTGAGTATTCGCGGCTTACCGGATGCGAGTGTGCAAGTGAATTCCTTTGAA 640
 QY 541 ATTAAGGCGCTGAGTGTGCGGCTTCTCTCTTGTGACGCTGACTTGTGATTCCTGC 600
 Db 641 ATTAAGGCGCTGAGTGTGCGGCTTCTCTCTTGTGACGCTGACTTGTGATTCCTGC 700
 QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCACTTCTCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCACTTCTCTT 760
 QY 661 GTGGTAATTCAGAGTACAGGCGCCCTATTTGGCGCGCTGCGATCTCTTGGGCTGTGAC 720
 Db 761 GTGGTAATTCAGAGTACAGGCGCCCTATTTGGCGCGCTGCGATCTCTTGGGCTGTGAC 820
 QY 721 ATCCGTAATCTCTTCTTGGGAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 780
 Db 821 ATCCGTAATCTCTTCTTGGGAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 880
 QY 781 TTGTTGTCGAGTATGTCAGTATCTTTTGGGCTCCGGCGGCTCCGTTCTTATCC 840
 Db 881 TTGTTGTCGAGTATGTCAGTATCTTTTGGGCTCCGGCGGCTCCGTTCTTATCC 940
 QY 841 TTAAGCCCTACGTGAATCAATTTGTGGCAAAATGGCGATGTGATGCGACAGAGAA 900
 Db 941 TTAAGCCCTACGTGAATCAATTTGTGGCAAAATGGCGATGTGATGCGACAGAGAA 1000

/transl_table=1
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="REMBL:CAC39923"
 /translation="MKTQELHSLSEVPSRALEPPDKGRYREINAGKTSIAGL
 GMYPIGIFGLVLIQYGEWMAAPLPSGLIFAGSTEMVLAVLVGAAPLGAIALTTLL
 NVFRVFAFSPFLVNVNPIARFYSVALIDEAVAVTAAPAPAGSAWRLISQIASH
 SYVVGGLGVIAELIPEIKGLEFALCSLFTVLTLDSCRTKKQIPSLIAGSFTI
 ALVVPQALPAAALIFGLLTIRYFGRKAK"
 GLVNF"

QY 901 TCCTGGCAATTTGACGATCAAGTTTGCAGCAATGCGATAGATCTGAAGCTCTAA 960
DB 1001 TCCTGGCAATTTGACGATCAAGTTTGCAGCAATGCGATAGATCTGAAGCTCTAA 1060
QY 961 CCTTTGGTCTATGCGCTTGGCATTTACAGTGTGGGCGCATCTTTGGGCGGTGAGCA 1020
DB 1061 CCTTTGGTCTATGCGCTTGGCATTTACAGTGTGGGCGCATCTTTGGGCGGTGAGCA 1120
QY 1021 CCTTTGGTGGGCGCTGCGATGATCTTTTGTGTGAGCTGGTGAATCTTTTC 1076
DB 1121 CCTTTGGTGGGCGCTGCGATGATCTTTTGTGTGAGCTGGTGAATCTTTTC 1176

RESULT 4
AX137714 1271 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 6 from Patent EP1096010.
DEFINITION AX137714
ACCESSION AX137714.1 GI:14273893
VERSION
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS
Kemernecht, N., Esseling, L., Salm, H. and Pfeifferle, W.
TITLE
Nucleotide sequences coding for branched-chain amino acids export
protein, method for isolating them and their use
Patent: EP 1096010-A 6 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
JOURNAL
Location/Qualifiers
FEATURES
source
1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC13032"
101. 856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=11
/protein_id="CAC39927.1"
/db_xref="GI:14273894"
/db_xref="EMBL:CAC39927"
/translation="MOQTQEIHSLSLEVPKSALEPDDKGRRYEIAQGLTSLAAGL
GVYPIAGFLVIOGYEMWAAPIFSGLTAGSTEMVILVVGAAPIGLATITLL
VNEHVFAPSPFLVKNPPIARFVPAIDBAVAVTAARPAQMSAMRLISQIAFH
SYWVGGATGVAAIELIPETIKGEPALCSLPTLTLDSCRTKQIFSLLAGSFTI
ALVVIPEGALFPAALITFLGLTIRIFLGRKAK"
853. 1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=11
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="EMBL:CAC39928"
/translation="WTDPSCLILVAACVITPALRAVPLILKPIRESQVGMAM
WMPAGILITLTASTRSNAIDKLTFLGLINAVITVAHLIGRRLLISVAGITVY
GLWLPF"

ORIGIN
Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275; Indels 0; Gaps 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTTCGCCATCCAGGACAGCCCTG 60
DB 101 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTTCGCCATCCAGGACAGCCCTG 160
QY 61 GAACCAAGATGAAGGTATGCGGCTTCGAAATGCGGACGATCTAAATCCCTCCCT 120
DB 161 GAACCAAGATGAAGGTATGCGGCTTCGAAATGCGGACGATCTAAATCCCTCCCT 220
QY 121 GCTGCAAGTTTGGGCGATGACCCGATGGATTCGTTGATCTCTGGTTATTCATATC 180

DB 221 GCTGCAAGTTTGGGCGATGACCCGATGGTATGTTGGTTGGTCTCTGGTATTCATATC 280
QY 181 GGCTACGAATGGTGGGACGCCCCCACTGTTTCCGGCTGATTTTGCAGGCTCCACCGAA 240
DB 281 GGCTACGAATGGTGGGACGCCCCCACTGTTTCCGGCTGATTTTGCAGGCTCCACCGAA 340
QY 241 ATGCTGGTATGCGCTTGGTGGGCGAGCGCCCTGCGGCGCATGCGCTCACCA 300
DB 341 ATGCTGGTATGCGCTTGGTGGGCGAGCGCCCTGCGGCGCATGCGCTCACCA 400
QY 301 TTGCTGGTGAAGTTCGCGAGATTTCTATGAGTTTCAATTCGCGTGAATGGTCAAA 360
DB 401 TTGCTGGTGAAGTTCGCGAGATTTCTATGAGTTTCAATTCGCGTGAATGGTCAAA 460
QY 361 AACCCCATGGCCCGTTCTATTCGCTTTTCGCGCTTATCGAAGCCGACGACTACT 420
DB 461 AACCCCATGGCCCGTTCTATTCGCTTTTCGCGCTTATCGAAGCCGACGACTACT 520
QY 421 GCGGCGAGCGCCGCGAGCTGCTGCGGCGTGGCGCATTTATCTCAATGCAATAGCGTTTAC 480
DB 521 GCGGCGAGCGCCGCGAGCTGCTGCGGCGTGGCGCATTTATCTCAATGCAATAGCGTTTAC 580
QY 481 TCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGGAGTGCAGAGTTGATTCCTTTGAA 540
DB 581 TCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGGAGTGCAGAGTTGATTCCTTTGAA 640
QY 541 ATTAAGGCGCTGAGTTCGCGCTTGTCTTCTTGTGTGACCTGACTTGGATTCCTTC 600
DB 641 ATTAAGGCGCTGAGTTCGCGCTTGTCTTCTTGTGTGACCTGACTTGGATTCCTTC 700
QY 601 CGAAGCAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGTTCACATTCCTCTT 660
DB 701 CGAAGCAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGTTCACATTCCTCTT 760
QY 661 GTGTGATTCGAGTCAAGCGCTTATTCGCGGCTGCTGATCTTCTGGGTCTGTTGACC 720
DB 761 GTGTGATTCGAGTCAAGCGCTTATTCGCGGCTGCTGATCTTCTGGGTCTGTTGACC 820
QY 721 ATCCGGTACTTCTTCTGGGAAAGGCTGTAAATGCAACTGATTCCTCTGATTCCTCC 780
DB 821 ATCCGGTACTTCTTCTGGGAAAGGCTGTAAATGCAACTGATTCCTCTGATTCCTCC 880
QY 781 TTGTGTGCGAGTATGAGTCAATTAATCTTTGCGCTCCGCGCGTTCCGTTCTTAATCC 840
DB 881 TTGTGTGCGAGTATGAGTCAATTAATCTTTGCGCTCCGCGCGTTCCGTTCTTAATCC 940
QY 841 TTAAAGCCCTCAGTGAATGACAAATTTGTGGGCAAAATGCGAATGTGATGCCAGAGAA 900
DB 941 TTAAAGCCCTCAGTGAATGACAAATTTGTGGGCAAAATGCGAATGTGATGCCAGAGAA 1000
QY 901 TCCTTGGCAATTTGACCGCATCAACGTTTTCGACGAATGCAATGATCTGAAGCTCTAA 960
DB 1001 TCCTTGGCAATTTGACCGCATCAACGTTTTCGACGAATGCAATGATCTGAAGCTCTAA 1060
QY 961 CCTTTGGTCTATGCGCTTGGCATTTACAGTGTGGGCGCATCTTTGGGCGGTGAGCA 1020
DB 1061 CCTTTGGTCTATGCGCTTGGCATTTACAGTGTGGGCGCATCTTTGGGCGGTGAGCA 1120
QY 1021 CCTTTGGTGGGCGCTGCGATGATCTTTTGTGTGAGCTGGTGAATCTTTTC 1076
DB 1121 CCTTTGGTGGGCGCTGCGATGATCTTTTGTGTGAGCTGGTGAATCTTTTC 1176

RESULT 5
BD014991 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014991
DEFINITION
Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION BD014991.1 GI:22555798
VERSION BD014991.1
KEYWORDS UP 2001189788-A/1.
SOURCE unidentified

ORGANISM unidentified
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kennerkumehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL JP 2001169788-A 1 26-JUN-2001;
DEUSHA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/1
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KENNERKUMEHITO, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
PC C12N15/00,
PC C12N15/00, C12R1:15)
CC bnf
CC bnf
CC bnf
FH key Location/Qualifiers
FT gene (101)..(853)
FT gene (853)..(1176).
FT gene Location/Qualifiers
1..1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTATTCAGACCTGAGAGGTGTGGCCATCCAAAGCAAGCCCTG 60
DB 101 GTGCAAAAACGCAAGATTATTCAGACCTGAGAGGTGTGGCCATCCAAAGCAAGCCCTG 160
QY 61 GAAACAGATGATTAAGTTATCGGCGCTACGAATCGCGCAAGGTCTAAAAACCTCCCTT 120
DB 161 GAAACAGATGATTAAGTTATCGGCGCTACGAATCGCGCAAGGTCTAAAAACCTCCCTT 220
QY 121 GCTGAGGTTTGGGAGATGATACCCGATGATGCTTGTCTCTTGGTTATTCATATC 180
DB 221 GCTGAGGTTTGGGAGATGATACCCGATGATGCTTGTCTCTTGGTTATTCATATC 280
QY 181 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAA 240
DB 281 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAA 340
QY 241 ATGCTGATGATGACCTGCTTGTGGGCGAGCGCCCTTGGGCGCATGCGCTCACCA 300
DB 341 ATGCTGATGATGACCTGCTTGTGGGCGAGCGCCCTTGGGCGCATGCGCTCACCA 400
QY 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
QY 361 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 461 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
QY 421 GGGGCGAGGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 521 GGGGCGAGGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
QY 481 TCTTACTGGGATATGCGGAGTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGA 540
DB 581 TCTTACTGGGATATGCGGAGTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGA 640
QY 541 ATTAAAGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

DB 641 ATTAAAGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
QY 601 CGAAGCAAAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 701 CGAAGCAAAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
QY 661 GTGATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 761 GTGATATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 820
QY 721 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 821 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 880
QY 781 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 881 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 940
QY 841 TTAAGCCCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 941 TTAAGCCCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
QY 901 TCTTGGCATTTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1001 TCTTGGCATTTTGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1060
QY 961 CTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1061 CTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120
QY 1021 CTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
DB 1121 CTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1176

RESULT 6
BD014994 1271 bp DNA linear PART 27-AUG-2002
LOCUS
DEFINITION
Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION
BD014994
VERSION
BD014994.1 GI:22555801
KEYWORDS
JP 2001169788-A/4.
SOURCE
unidentified
ORGANISM
unidentified
REFERENCE
1 (bases 1 to 1271)
AUTHORS Kennerkumehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL JP 2001169788-A 1 26-JUN-2001;
DEUSHA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC13032
PN JP 2001169788-A/4
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KENNERKUMEHITO, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
PC C12N15/00,
PC C12N15/00, C12R1:15)
CC bnf
CC bnf
CC bnf
FH key Location/Qualifiers
FT gene (101)..(853)
FT gene (853)..(1176).
FT gene Location/Qualifiers
1..1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

/product="Lrp-like regulator"
 /protein_id="AA046687.1"
 /db_xref="GI:21311382"
 /translation="MKLDSIDRAIIAEISANARISNIALADKYLTPGPCRRVORBLE
 AEGIIIGASADIDHPANVRGEFVVDVTLNPSFVSADGDEVLAEHLRFLG
 SPDFVRIAGVADLEAYEQLSSHQITPGIAKISSRPMKVVVAPARQV"

ORIGIN

Query Match 100.0%; Score 1076; DB 1; Length 2105;
 Best local Similarity 100.0%; Pred. No. 8.8e-275;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAAGCAGAGATTCATTCAAGCCTGAGAGGCTGCATCCAGAGCCCTG 60
 Db 1463 GTGCAAAAAAGCAGAGATTCATTCAAGCCTGAGAGGCTGCATCCAGAGCCCTG 1404
 Qy 61 GAACCAAGATGATAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAACTCCCTT 120
 Db 1403 GAACCAAGATGATAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAACTCCCTT 1344
 Qy 121 GCTGCAAGTTGGGCAATGATACCGCATGATATGCGCTTTGGTCTCTGTTATTCATAC 180
 Db 1243 GCTGCAAGTTGGGCAATGATACCGCATGATATGCGCTTTGGTCTCTGTTATTCATAC 1284
 Qy 181 GGCTACGATGATGAGGAGAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
 Db 1283 GGCTACGATGATGAGGAGAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 1224
 Qy 241 ATGCTGATATGCGCTCTGTTGTTGGGCGAGCGCCCTGAGCGCATGCGCTACACACA 300
 Db 1223 ATGCTGATATGCGCTCTGTTGTTGGGCGAGCGCCCTGAGCGCATGCGCTACACACA 1164
 Qy 301 TTGCTGATGATCTCCGCAAGTATTCATGATGATTTTCAATCCCGCTGATGATGATGATA 360
 Db 1163 TTGCTGATGATCTCCGCAAGTATTCATGATGATTTTCAATCCCGCTGATGATGATGATA 1104
 Qy 361 AACCCATGCGCGCTTTTATTCGCTTTGCGCTTATGACGAAAGCTTACGAGTCACT 420
 Db 1103 AACCCATGCGCGCTTTTATTCGCTTTGCGCTTATGACGAAAGCTTACGAGTCACT 1044
 Qy 421 GGGGCAAGCCCGGAGGCTGATGCGCGCGGAGCTTATTCATGCAATGCAATGAGCTTTCAC 480
 Db 1043 GGGGCAAGCCCGGAGGCTGATGCGCGCGGAGCTTATTCATGCAATGCAATGAGCTTTCAC 984
 Qy 481 TCCATCGGATATTCGCGGATCTCAACGAGTGGCGATCGAGAGTTGATTCCTTTTGA 540
 Db 983 TCCATCGGATATTCGCGGATCTCAACGAGTGGCGATCGAGAGTTGATTCCTTTTGA 924
 Qy 541 ATTAAAGGCTCGAGTTGCGCTTGTCTCTTTTTCAGCGCTGATGATTTGATTCCTG 600
 Db 923 ATTAAAGGCTCGAGTTGCGCTTGTCTCTTTTTCAGCGCTGATGATTTGATTCCTG 864
 Qy 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTACCTTCAACCATGCTCTT 660
 Db 863 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTACCTTCAACCATGCTCTT 804
 Qy 661 GTGGAATTCAGGTCAGGCTTATTCGCGCGCTGCTGATCTTCTTGGGCTCTTGAAC 720
 Db 803 GTGGAATTCAGGTCAGGCTTATTCGCGCGCTGCTGATCTTCTTGGGCTCTTGAAC 744
 Qy 721 ATCCGATCTCTTTTGGGAAAAGGCTGCTAAATGACAAGTATTCCTCGTATTCCTC 780
 Db 743 ATCCGATCTCTTTTGGGAAAAGGCTGCTAAATGACAAGTATTCCTCGTATTCCTC 684
 Qy 781 TTGTTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 Db 683 TTGTTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 624
 Qy 841 TTAAGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 Db 623 TTAAGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 564
 Qy 901 TCCCTTGCATTTTGAACGATCAACGTTTTCGAGCAATGCGATGATGATGATGATGATGAT 960

RESULT 8
 AP005274
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

563 TCCCTTGCATTTTGAACCGGATCAAGCTTTCGAGAAAGCGATATCTGAAGACTCTA 504
 Qy 961 CTTTGTGCTCATTCACCGCTTGCATTAACAGTGTGCGCATCTTGTGCGGTGACGCA 1020
 Db 503 CTTTGTGCTCATTCACCGCTTGCATTAACAGTGTGCGCATCTTGTGCGGTGACGCA 444
 Qy 1021 CTTTGTGAGCGTTGGCGCTGCGCACCATGCTTTTGTGATGATGATGATGATGATGAT 1076
 Db 443 CTTTGTGAGCGTTGGCGCTGCGCACCATGCTTTTGTGATGATGATGATGATGATGAT 388

AP005274 340000 bp DNA linear BCT 08-AUG-2002
 Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
 AP005274 BA000036
 AP005274.1 GI:21322764
 Corynebacterium glutamicum ATCC 13032
 Corynebacterium glutamicum ATCC 13032
 Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
 1
 Nakagawa, S.
 Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
 Unpublished
 2 (bases 1 to 340000)
 Nakagawa, S.
 Direct Submission
 Submitted (24-MAY-2002) Satochi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Aabhi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@kankogen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
 This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.
 Location/Qualifiers
 1..340000
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 1..1575
 /gene="Cg10001"
 1..1575
 /gene="Cg10001"
 /note="PB00308: Bacterial dna protein
 TIGR00362: DnaA: chromosomal replication initiator protein
 DnaA"
 /codon_start=1
 /transl_table=1
 /product="ATPase involved in DNA replication initiation"
 /protein_id="BAB97394.1"
 /db_xref="GI:21322765"
 /translation="MSQNSSLLETRVAVADLTLSQADSGFPIPTORAVNLT
 KPIALVDGAVLSTFVAKAVYENDLGLALTRVLSKRGSPSAVSEPOEIPET
 PAQOEKYPDAVVISNPAKPYEVGGGASTDGMERTSAPPEHPAPVADPE
 PELATPQRI.PRETPANPRVSLNPKTFESFVIGPFRFANAANAVAASPARAFN
 PLFISGSGSLGTHLHAGVAGOEPLRIKYVSEFTNDYVNSVDDOEFK
 RKNLDILMVDIOFLAGKEGTQEEFTFNLHQAOKOILSDRPKQLTLEDL
 RRPREGGLTIDOPLETERRIAIIMKKQTDGTHVDRVLEIARFESISELGA
 IYSAYSLINQPIKEMAIVLRDILPEPMEITAPVIMETVAPEPISVDTLGA
 GTRVAYAHKQAMTICRELITMSLPKIGDVGGKDHITVTADKRIQEMERKDTY
 DETQUTQILKSRGRN"
 /complement (1594..1920)
 /gene="Cg10002"
 /complement (1594..1920)
 /gene="Cg10002"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"


```

gene
CDS
    /protein_id="BAB97395.1"
    /db_xref="GI:21322766"
    /translation="MIVNPFVLMIDFKRLIPVCADVPAYPOFLGSCAQGKFRG
    ITPPTKLTQLQVNCYTAIVKVNDFVSMAFISYQNPCTTPTKIQCPBDSGSSRA
    ELNPF1"
    2292..3476
    /gene="Cg10003"
    2292..3476
    /gene="Cg10003"
    /EC_number="2.7.7.7"
    /note="PF00712: DNA polymerase III beta subunit, N-terminal
    domain"
    PF02767: DNA polymerase III beta subunit, central domain
    PF02768: DNA polymerase III beta subunit, C-terminal domain
    TIGR00663: dhan: DNA polymerase III, beta subunit"
    /codon_start=1
    /transl_table=11
    /product="DNA polymerase III beta subunit"
    /protein_id="BAB97396.1"
    /db_xref="GI:21322767"
    /translation="MESQVSTFRVAREDLVTAVAVANLPTKPTQVPLAMLTADD
    EQLBLAGYDVSTVRLSAEVSQGRIVAGKLSEITGSLPNKPVDFRIDSKAFV
    TCGSRFELPLIPDDYPMPLPAGTGINAKLFTBAVSQVSAAGKDSLEMLTGV
    SNEIVGNCINLATRFRALPTFEMEANDEPLNKLIPARTLIDNARNSDGLNS
    IDIAYGTQVQVREGFGVYTDNRRTTRLDADPNITAPLPEKHTAIASVETAPV
    DAIRVSLVAERNQIVLAFSRGVITLTAAGTAAGAAETTPCAFTQKELTIANPGY
    LKDGSLVPTSAVAFGFTSPRPAIMIPPEPMPSANENGIFQTPDITFTYLLNPLV
    PL"
    3585..4769
    /gene="Cg10004"
    3585..4769
    /gene="Cg10004"
    /note="TIGR00611: recf: Recf protein"
    /codon_start=1
    /transl_table=11
    /product="Recombinational DNA repair APase"
    /protein_id="BAB97397.1"
    /db_xref="GI:21322768"
    /translation="MIRSELEDRYSMPLEKVDDEPGITVFIGNGFGKTNIEAIG
    VLAHSRHSVSDAPLVRHAENARSAVANQRELAHLIKPHANQSNLRTKY
    RPRRLGIVYKTVLEAPDIALVKGPRARPRYLDITATQPRPAAGKAYDXTKY
    ENALTKRTIARRRGTEEGAAALSTIDGQCLARLGAEMARAPRIBELGKTY
    KNYLTAPSRPAANVYKTTIDQGSQSEFPAGILIEATLLETAKORBEIGSS
    VQPHRDDVLMQDPAGKGFASGHTSFALSLRIABENLTKSDTDIILLIDVFS
    LDAGREKLVGIAQVEQVLTITAAVHDLPEMLKVTIQAHTVIVQDTGTRISILDV
    QP"
    4766..5302
    /gene="Cg10005"
    4766..5302
    /gene="Cg10005"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB97398.1"
    /db_xref="GI:21322769"
    /translation="MTDPIEQAFERIRAEAMRANGSVPLDNKDAFRPPAKGVK
    RKKGPSGIDGQKIVRGASLVINKETQRGNGKQIRAGWVTSWEELVGAKIA
    QHTREVMKDKLFTTCDSTAWATNLRMWQRIQIVIAEKVGNIIITELRFGPAPS
    MKKGPLHVGRSPRDYTG"
    5354..7489
    /gene="Cg10006"
    5354..7489
    /gene="Cg10006"
    /EC_number="5.99.1.3"
    /note="PF00204: DNA topoisomerase II (N-terminal region)
    PF00986: DNA gyrase B subunit, carboxyl terminus
    PF01751: Toprim domain
    PF02518: Histidine kinase, DNA gyrase B-, phytochrome-like
    ATPase
    TIGR01055: pare: GneG: DNA topoisomerase IV, B subunit
    TIGR01058: pare: Gpos: DNA topoisomerase IV, B subunit
    TIGR01059: gyrB: DNA gyrase, B subunit"
    /codon_start=1

```

```

gene
CDS
    /transl_table=11
    /product="DNA gyrase (topoisomerase II) B subunit"
    /protein_id="BAB97399.1"
    /db_xref="GI:21322770"
    /translation="MRGTTWGPKRVRMKRLYRISSEBGLKANTENHYDASSITILE
    GLEAVRBRQVYSGPRGLHLIKREVDNSVDAMAGAKVTVTLLEDGVOYV
    DNRGITVDNHPGSAFTVQVYVQTLHAGGFRDSDYAVGSLHGVGCSYVNALSTVEA
    DKILHGRHWYQNFESVPDELTIEGNARGTGTTIRFVPDAIIFETIEPDETTISRLQ
    ENAFILNKGLITLTDNRATDEBELALAEQELVLESLDIEDNETVEETTDAPK
    KPKRREKKKI FHYPNGLEDYVHYLNRSKTNIPSTIVSFEAKGDHDEVEVAMQNSYK
    ESVHTFANTINREGSTHEGFRSALTSIMNRVAREKLLKEKENALTGDDCREGLSA
    VISVRVDDPQFESQTKLGNTEIKSFVQRVNRNEHGHLEANPAEAKYTIKAVGSA
    QRLAARADIVRSKATDIDGLPGKLAQVQVITRLGTHDEPDKRHKRYLM
    SMQALPLRGLKILNREKRLDKYLNKAEVQAITRLGTHDEPDKRHKRYLM
    ADADVQGHATILTLTLRFMPDLVAEGVTLAQPPLKLMQKQEBQFQFAYSDBERD
    EQLNEGLAAGRKINKDDGIQRKGLGENNAELMETWMDPTVIRLRVSDITDAQRDE
    LFSILMGDDVVARSRFTIRNAKDVRLDT"
    7830..8798
    /gene="Cg10007"
    7830..8798
    /gene="Cg10007"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB97400.1"
    /db_xref="GI:21322771"
    /translation="MTCYTIKIOSIKNAQREMLKTTLTITDPLGHVTSPEPTTHAVP
    LPDGSSTPVOIYASDNKDSQVLMNPFGRMGCTYTRPLAANKAGFHAIGELAGQ
    OSSAKARSKSQWGHDLASVDFPQIAAKKALDDEBGMPLSHSKGQSLSCLPFA
    RPEAKRYNLRAIFGVAGSGPFRFTSPKXGKLGAVALGGIIGHIIVGFPGKVLGK
    DLVGYRSGSTHREMRFRHKNSLDLDTAQDINVEVWKVSIPTTFRCRCDDBCP
    QASIDGLASFVPAQIKMIEIPALGHNWAREPESTVLPLEQAL"
    complement(8795..9466)
    /gene="Cg10008"
    complement(8795..9466)
    /gene="Cg10008"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical membrane protein"
    /protein_id="BAB97401.1"
    /db_xref="GI:21322772"
    /translation="KGNPNKKPDLPAQRESIPRNAMFVFGIFIMSFAIGTVHAGIG
    TTITSLPVMYTAGSLGSLGWTTITFNGNMLIQCIIVLSQKRPQVLIILNAPLFGP
    LCOLSLQLTWAGTDNIVFVAMINVISITLMSIGVITQVLPNITFAGSIVSALVK
    FPNVERGTMKQIVMTFVSAALISITWGLIGVNEGTVFAAFIFGFPVRQMKLYL
    RSLIGH"
    complement(9471..9914)
    /gene="Cg10009"
    complement(9471..9914)
    /gene="Cg10009"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB97402.1"
    /db_xref="GI:21322773"
    /translation="MTSLDILVYSTADAADMSAFIRETVGKSGISGIALKGAMG
    AATKISDIIYAKSRKLLPEIADSLDGLNOBYONGSTADPQHLETPANSSTALDALS
    VADKNASINIVPELGVYVGRGKAANKVLEDELPAIGLIEKNAQ"
    complement(10104..11177)
    /gene="Cg10010"
    complement(10104..11177)
    /gene="Cg10010"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB97403.1"
    /db_xref="GI:21322774"
    /translation="MFMIAQRITPIHITAPHLFVARVQIRATDADRSLGRLIELS

```

Query Match 100.0%; Score 1076; DB 1; Length 340000;
 Best Local Similarity 100.0%; Pred. No. 1.5e-274; Indels 0; Gaps 0;
 Matches 1076; Conservative 0; Mismatches 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 60
 Db 276829 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 276888
 QY 61 GAACCAAGATGATTAAGGTTATCGGGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCTT 120
 Db 276889 GAACCAAGATGATTAAGGTTATCGGGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCTT 276948
 QY 121 GCTGCAAGGTTTGGGCAATGATCCGATGGAATGGAATGCGTTTGTCTTGGTTATTCAATAC 180
 Db 276949 GCTGCAAGGTTTGGGCAATGATCCGATGGAATGGAATGCGTTTGTCTTGGTTATTCAATAC 277008
 QY 181 GGCTACGAATGTTGGGCAAGCCCACTGTTTCCGGCCCGATTTTCGGGGCTCCACCGAA 240
 Db 277009 GGCTACGAATGTTGGGCAAGCCCACTGTTTCCGGCCCGATTTTCGGGGCTCCACCGAA 277068
 QY 241 ATGCTGATATCGCCCTGTTGTGGGCGCAAGCGCCCTGAGGCCATCGCGCTCACCA 300
 Db 277069 ATGCTGATATCGCCCTGTTGTGGGCGCAAGCGCCCTGAGGCCATCGCGCTCACCA 277128
 QY 301 TTGCTGTGAATCTCCGCAAGTATTCTATGCGTTTTCATTCGGCTGCATGTGCTCAA 360
 Db 277129 TTGCTGTGAATCTCCGCAAGTATTCTATGCGTTTTCATTCGGCTGCATGTGCTCAA 277188
 QY 361 AACCCATTGCGCCGTTCTATTCGGTTTTCGGCTTATCGACGAAGCCTACGCACTACT 420
 Db 277189 AACCCATTGCGCCGTTCTATTCGGTTTTCGGCTTATCGACGAAGCCTACGCACTACT 277248
 QY 421 GCGGCGAGGCGCGAGGCTGTGCGGCGGCGCACTTATCTCAATGCAAAATGCGTTTAC 480
 Db 277249 GCGGCGAGGCGCGAGGCTGTGCGGCGGCGCACTTATCTCAATGCAAAATGCGTTTAC 277308
 QY 481 TCCCTAGGATATTCGGGCGGTCTCACGGAAGTGGCGATGCAAGATTGATTCCTTTGAA 540
 Db 277309 TCCCTAGGATATTCGGGCGGTCTCACGGAAGTGGCGATGCAAGATTGATTCCTTTGAA 277368
 QY 541 ATTAAGGACCTGAGTTGCGCCTTGTGCTCTCTTGTGACGCTGACTTGGATTCCTGC 600
 Db 277369 ATTAAGGACCTGAGTTGCGCCTTGTGCTCTCTTGTGACGCTGACTTGGATTCCTGC 277428
 QY 601 CGAAGCAAAAGCAAGATCCCTTCTGCTGCTGCGAGGTTTGAAGTTACCATGCTCTT 660
 Db 277429 CGAAGCAAAAGCAAGATCCCTTCTGCTGCTGCGAGGTTTGAAGTTACCATGCTCTT 277488
 QY 661 GTGTAATTCAGATCAGGACCTATTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 720
 Db 277489 GTGTAATTCAGATCAGGACCTATTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 277548
 QY 721 ATCCGATCTTCTTCTTGGGAAAGGCTGTAAATGACAACTGATTTCTCTGTATTCCTC 780
 Db 277549 ATCCGATCTTCTTCTTGGGAAAGGCTGTAAATGACAACTGATTTCTCTGTATTCCTC 277608
 QY 781 TTGTTGTGAGATGTCAGATCATTACTTTTGCCTCGGAGCGGTTCGTTCTTAATCC 840
 Db 277609 TTGTTGTGAGATGTCAGATCATTACTTTTGCCTCGGAGCGGTTCGTTCTTAATCC 277668
 QY 841 TTAAAGCCCTTACGTAATCAATTTGTGGGCAAAATGCGCAATGTGATGTCACAGGAA 900
 Db 277669 TTAAAGCCCTTACGTAATCAATTTGTGGGCAAAATGCGCAATGTGATGTCACAGGAA 277728
 QY 901 TCCCTGCAATTTTACGCGATCAAGTTTTCGAGCAATGCAATGATGTAAGAATCTTAA 960
 Db 277729 TCCCTGCAATTTTACGCGATCAAGTTTTCGAGCAATGCAATGATGTAAGAATCTTAA 277788
 QY 961 CCTTTGTCTCATTTGCGGTTGATTAAGTGTGGCGCATCTTCTTGGCGGTGACGCA 1020
 Db 277789 CCTTTGTCTCATTTGCGGTTGATTAAGTGTGGCGCATCTTCTTGGCGGTGACGCA 277848
 QY 1021 CCTTTGTAGGTTTGGCGGTGCGCAATGCTTTTGTGAGTGTGATCTTTTTC 1076
 Db 277849 CCTTTGTAGGTTTGGCGGTGCGCAATGCTTTTGTGAGTGTGATCTTTTTC 277904

RESULT 9
 AX120085 349980 bp DNA linear PAT 11-MAY-2001
 LOCUS AX120085
 DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:114036800
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source
 1. 349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0.000.001 0.349.980 349.980"
 ORIGIN
 Query Match 100.0%; Score 1076; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 1.5e-274; Mismatches 0; Gaps 0;
 Matches 1076; Conservative 0; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 60
 Db 276829 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 276888
 QY 61 GAACCAAGATGATTAAGGTTATCGGGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCTT 120
 Db 276889 GAACCAAGATGATTAAGGTTATCGGGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCTT 276948
 QY 121 GCTGCAAGGTTTGGGCAATGATCCGATGGAATGGAATGCGTTTGTCTTGGTTATTCAATAC 180
 Db 277309 GCTGCAAGGTTTGGGCAATGATCCGATGGAATGGAATGCGTTTGTCTTGGTTATTCAATAC 277368
 QY 181 GGCTACGAATGTTGGGCAAGCCCACTGTTTCCGGCCCGATTTTCGGGGCTCCACCGAA 240
 Db 277429 GGCTACGAATGTTGGGCAAGCCCACTGTTTCCGGCCCGATTTTCGGGGCTCCACCGAA 277488
 QY 241 ATGCTGATATCGCCCTGTTGTGGGCGCAAGCGCCCTGAGGCCATCGCGCTCACCA 300
 Db 277069 ATGCTGATATCGCCCTGTTGTGGGCGCAAGCGCCCTGAGGCCATCGCGCTCACCA 277128
 QY 301 TTGCTGTGAATCTCCGCAAGTATTCTATGCGTTTTCATTCGGCTGCATGTGCTCAA 360
 Db 277129 TTGCTGTGAATCTCCGCAAGTATTCTATGCGTTTTCATTCGGCTGCATGTGCTCAA 277188
 QY 361 AACCCATTGCGCCGTTCTATTCGGTTTTCGGCTTATCGACGAAGCCTACGCACTACT 420
 Db 277189 AACCCATTGCGCCGTTCTATTCGGTTTTCGGCTTATCGACGAAGCCTACGCACTACT 277248
 QY 421 GCGGCGAGGCGCGAGGCTGTGCGGCGGCGCACTTATCTCAATGCAAAATGCGTTTAC 480
 Db 277249 GCGGCGAGGCGCGAGGCTGTGCGGCGGCGCACTTATCTCAATGCAAAATGCGTTTAC 277308
 QY 481 TCCCTAGGATATTCGGGCGGTCTCACGGAAGTGGCGATGCAAGATTGATTCCTTTGAA 540
 Db 277309 TCCCTAGGATATTCGGGCGGTCTCACGGAAGTGGCGATGCAAGATTGATTCCTTTGAA 277368
 QY 541 ATTAAGGACCTGAGTTGCGCCTTGTGCTCTCTTGTGACGCTGACTTGGATTCCTGC 600
 Db 277369 ATTAAGGACCTGAGTTGCGCCTTGTGCTCTCTTGTGACGCTGACTTGGATTCCTGC 277428
 QY 601 CGAAGCAAAAGCAAGATCCCTTCTGCTGCTGCGAGGTTTGAAGTTACCATGCTCTT 660
 Db 277429 CGAAGCAAAAGCAAGATCCCTTCTGCTGCTGCGAGGTTTGAAGTTACCATGCTCTT 277488

QY 661 GTGTAATTCAGGTCAGGCGCTATTTCGCGGCTGATGATCTTCTGGGTCTGTGACC 720
 DB 277489 GTGGTAATTCAGGTCAGGCGCTATTTCGCGGCTGATGATCTTCTGGGTCTGTGACC 277548
 QY 721 ATCCGGTACTTCTTCTGGGAAAGGCTGTCTAATGACACTGATTTCTCTGATATTCCTC 780
 DB 277549 ATCCGGTACTTCTTCTGGGAAAGGCTGTCTAATGACACTGATTTCTCTGATATTCCTC 277608
 QY 781 TTGTTGCGCAGTATGTCAGTCACTTCTTGGGCTCCGGGCGGTTCCGTTCTTAATCC 840
 DB 277609 TTGTTGCGCAGTATGTCAGTCACTTCTTGGGCTCCGGGCGGTTCCGTTCTTAATCC 277668
 QY 841 TTAAGCCCTTACGTGAATCACAATTTGTGGGCAAAATGGCGATGTGATGCGCAGAGAA 900
 DB 277669 TTAAGCCCTTACGTGAATCACAATTTGTGGGCAAAATGGCGATGTGATGCGCAGAGAA 277728
 QY 901 TCCCTGGCATTTTGAACCGCATCAAGCTTTCGAGCAATGCGATGATGTAAGACTCTAA 960
 DB 277729 TCCCTGGCATTTTGAACCGCATCAAGCTTTCGAGCAATGCGATGATGTAAGACTCTAA 277788
 QY 961 CCTTGGTCTCATTCGCGTGGCATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 1020
 DB 277789 CCTTGGTCTCATTCGCGTGGCATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 277848
 QY 1021 CCTTGGTCTCATTCGCGTGGCATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 1076
 DB 277849 CCTTGGTCTCATTCGCGTGGCATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 277904

RESULT 10
 LOCUS AR391954 753 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 2 from patent US 6613545.
 ACCESSION AR391954
 VERSION AR391954.1 GI:40115725
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 753)
 Kemerhnecht,N., Salm,H., Eggeling,L. and Pfefferle,W.
 Nucleotide sequences coding for the export of branched amino
 acids, process for the isolation thereof and use thereof
 JOURNAL Patent: US 6613545-A 2 02-SEP-2003;
 FEATURES
 source 1..753
 /organism="Unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 70.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 5.6e-189;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACCGAAGATTCATTCAGGCTGAGGTGTGCCCATCCAGGACCCCTG 60
 DB 1 GTGCAAAAACCGAAGATTCATTCAGGCTGAGGTGTGCCCATCCAGGACCCCTG 60
 QY 61 GAACGATGATTAAGGTTATCGGGCTACGAATTCGCGCAAGTCTAAAAACCTCCCT 120
 DB 61 GAACGATGATTAAGGTTATCGGGCTACGAATTCGCGCAAGTCTAAAAACCTCCCT 120
 QY 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTGTCTCTGTTATTCATAC 180
 DB 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTGTCTCTGTTATTCATAC 180
 QY 181 GGCTACGATGATGAGGCGACCGCACTGTTTCGCGCTGATTTTGGCGGGCTCCACGAA 240
 DB 181 GGCTACGATGATGAGGCGACCGCACTGTTTCGCGCTGATTTTGGCGGGCTCCACGAA 240
 QY 241 ATGCTGATCATGCGCTGTTGTGGGCGACGCGCCCTGGGCGCATGCGCTCAACCA 300
 DB 241 ATGCTGATCATGCGCTGTTGTGGGCGACGCGCCCTGGGCGCATGCGCTCAACCA 300

DB 241 ATGCTGATCATGCGCTGTTGTGGGCGACGCGCCCTGGGCGCATGCGCTCAACCA 300
 QY 301 TTGCTGTGAACCTTCCGACAGTATTCATGAGTTTCAATCCGCTGATGTGCTAA 360
 DB 301 TTGCTGTGAACCTTCCGACAGTATTCATGAGTTTCAATCCGCTGATGTGCTAA 360
 QY 361 AACCCATTGCGCGTTCTTATGCGTTTTCGCGCTTATGACGAAGCTTACGACGACT 420
 DB 361 AACCCATTGCGCGTTCTTATGCGTTTTCGCGCTTATGACGAAGCTTACGACGACT 420
 QY 421 GCGGCGAGCGCGACAGTGTGCGGCGGCACTTATCTAATGAATAGCGTTTAC 480
 DB 421 GCGGCGAGCGCGACAGTGTGCGGCGGCACTTATCTAATGAATAGCGTTTAC 480
 QY 481 TCCCTACTGGGATTTCCGCGGTCTCACCGAGTGGCATGCGAGAGTTGATCTTTTGA 540
 DB 481 TCCCTACTGGGATTTCCGCGGTCTCACCGAGTGGCATGCGAGAGTTGATCTTTTGA 540
 QY 541 ATTAAGGCGCTGAGTTGCGCCCTTGTGCTCTCTTGTGACGCTGACTTGTGATCTTGC 600
 DB 541 ATTAAGGCGCTGAGTTGCGCCCTTGTGCTCTCTTGTGACGCTGACTTGTGATCTTGC 600
 QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGTTTGAAGCTTCAACATTCCTT 660
 DB 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGTTTGAAGCTTCAACATTCCTT 660
 QY 661 GTGTAATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 DB 661 GTGTAATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 QY 721 ATCCGTAATTCCTTCTTGGGAAAGGCTGCTAA 753
 DB 721 ATCCGTAATTCCTTCTTGGGAAAGGCTGCTAA 753

RESULT 11
 LOCUS AX120373 753 bp DNA linear PAT 11-MAY-2001
 DEFINITION Sequence 289 from Patent EP1108790.
 ACCESSION AX120373
 VERSION AX120373.1 GI:14037088
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterinae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 Nakagawa,S., Mizoguchi,H., Ando,S., Hayaishi,M., Ochiai,K.,
 Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
 Novel polynucleotides
 Patent: EP 1108790-A 289 20-JUN-2001;
 JOURNAL KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source 1..753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN
 Query Match 70.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 5.6e-189;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACCGAAGATTCATTCAGGCTGAGGTGTGCCCATCCAGGACCCCTG 60
 DB 1 GTGCAAAAACCGAAGATTCATTCAGGCTGAGGTGTGCCCATCCAGGACCCCTG 60
 QY 61 GAACGATGATTAAGGTTATCGGGCTACGAATTCGCGCAAGTCTAAAAACCTCCCT 120
 DB 61 GAACGATGATTAAGGTTATCGGGCTACGAATTCGCGCAAGTCTAAAAACCTCCCT 120
 QY 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTGTCTCTGTTATTCATAC 180
 DB 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTGTCTCTGTTATTCATAC 180

Db 121 GCTGAGGTTGGGATGATACCCGATGGTATGCGTTGCTCTGTTATTCAATAC 180
 Qy 121 GGGTACGAATGGTGGGAGAGCCGCTGTTTCCGGCCGATTTCCGGGGCTCCACCGAA 240
 Db 121 GGGTACGAATGGTGGGAGAGCCGCTGTTTCCGGCCGATTTCCGGGGCTCCACCGAA 240
 Qy 241 ATGCTGATCATGCGCCCTGTTGAGGAGAGCCGCTGAGGCGCATCGCGTCAACCA 300
 Db 241 ATGCTGATCATGCGCCCTGTTGAGGAGAGCCGCTGAGGCGCATCGCGTCAACCA 300
 Qy 301 TTGCTGGTGAATCTTCCGCAAGTATTCATGAGTTTCAATCCCGCTGATGAGTCAA 360
 Db 301 TTGCTGGTGAATCTTCCGCAAGTATTCATGAGTTTCAATCCCGCTGATGAGTCAA 360
 Qy 361 AACCCATTGCGCGTTTCTAATCGGTTTCCGCTTATGAGAGCAAGCTGACAGTCACT 420
 Db 361 AACCCATTGCGCGTTTCTAATCGGTTTCCGCTTATGAGAGCAAGCTGACAGTCACT 420
 Qy 421 GCGGCCAGGCGCCGAGGCTGAGGCGGCTGAGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
 Db 421 GCGGCCAGGCGCCGAGGCTGAGGCGGCTGAGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
 Qy 481 TCCCTAGGATATGAGGAGGCTGAGGAGGAGTGGAGTGGAGTGAATCTTTGAA 540
 Db 481 TCCCTAGGATATGAGGAGGCTGAGGAGGAGTGGAGTGGAGTGAATCTTTGAA 540
 Qy 541 ATTAAGGAGCTGAGATGCGCCCTTTCCTCTCTCTTTCACGCTGATCTTGGATTCTGTC 600
 Db 541 ATTAAGGAGCTGAGATGCGCCCTTTCCTCTCTCTTTCACGCTGATCTTGGATTCTGTC 600
 Qy 601 CGAAGGAAAGAAAGATCCCTTCTGCTGCTGCGAGGTTGAGTTGACATTCATGCTCTT 660
 Db 601 CGAAGGAAAGAAAGATCCCTTCTGCTGCTGCGAGGTTGAGTTGACATTCATGCTCTT 660
 Qy 661 GTGTAATTCAGATGAGGCGCTATTTGAGGAGTGTGATCTTGGGTCTGTGACC 720
 Db 661 GTGTAATTCAGATGAGGCGCTATTTGAGGAGTGTGATCTTGGGTCTGTGACC 720
 Qy 721 ATCCGACTTCTTCTTGGGAAAGGCTGCTAA 753
 Db 721 ATCCGACTTCTTCTTGGGAAAGGCTGCTAA 753

RESULT 12

LOCUS AX137710 753 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 2 from Patent EP1096010.
 ACCESSION AX137710
 VERSION AX137710.1 GI:14273889
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteriia; Actinobacteriales; Actinomycetales;
 Corynebacteriaceae; Corynebacteriales; Corynebacterium.
 REFERENCE
 1 Kemernecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 2 02-MAY-2001;
 JOURNAL Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 FEATURES
 source
 1. 753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1. >753
 /note="unnamed protein product; brmf"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39925.1"
 /db_xref="GI:14273890"
 /db_xref="REFSEQ:CAC39925"

CDS

1. >753
 /note="unnamed protein product; brmf"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39925.1"
 /db_xref="GI:14273890"
 /db_xref="REFSEQ:CAC39925"

ORIGIN

Query Match 70.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 5 6e-189;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAAGGCAAGATTCATTAAGCTTGAAGGTGTGCTGATTCAGAGGAGCTTG 60
 Db 1 GTGCAAAAAGGCAAGATTCATTAAGCTTGAAGGTGTGCTGATTCAGAGGAGCTTG 60
 Qy 61 GAACCAATGATTAAGTTATCGGCGCTAGAAATGGCGCAAGGTCTAAAACTCCCTT 120
 Db 61 GAACCAATGATTAAGTTATCGGCGCTAGAAATGGCGCAAGGTCTAAAACTCCCTT 120
 Qy 121 GCTGCAAGTTGGGCAATGATCCGATGATGATGCTTGTGCTCTTGGTTATTCAATAC 180
 Db 121 GCTGCAAGTTGGGCAATGATCCGATGATGATGCTTGTGCTCTTGGTTATTCAATAC 180
 Qy 181 GGTACGAATGATGGGCAAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACGAA 240
 Db 181 GGTACGAATGATGGGCAAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACGAA 240
 Qy 241 ATGCTGATCATGCGCCCTGTTGAGGAGAGGCGCCCTGAGGCGCATCGCGTCAACCA 300
 Db 241 ATGCTGATCATGCGCCCTGTTGAGGAGAGGCGCCCTGAGGCGCATCGCGTCAACCA 300
 Qy 301 TTGCTGGTGAATCTTCCGCAAGTATTCATGAGTTTCAATCCCGCTGATGAGTCAA 360
 Db 301 TTGCTGGTGAATCTTCCGCAAGTATTCATGAGTTTCAATCCCGCTGATGAGTCAA 360
 Qy 361 AACCCATTGCGCGTTTCTAATCGGTTTCCGCTTATGAGAGCAAGCTGACAGTCACT 420
 Db 361 AACCCATTGCGCGTTTCTAATCGGTTTCCGCTTATGAGAGCAAGCTGACAGTCACT 420
 Qy 421 GCGGCCAGGCGCCGAGGCTGAGGCGGCTGAGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
 Db 421 GCGGCCAGGCGCCGAGGCTGAGGCGGCTGAGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
 Qy 481 TCCCTAGGATATGAGGAGGCTGAGGAGGAGTGGAGTGGAGTGAATCTTTGAA 540
 Db 481 TCCCTAGGATATGAGGAGGCTGAGGAGGAGTGGAGTGGAGTGAATCTTTGAA 540
 Qy 541 ATTAAGGAGCTGAGATGCGCCCTTTCCTCTCTCTTTCACGCTGATCTTGGATTCTGTC 600
 Db 541 ATTAAGGAGCTGAGATGCGCCCTTTCCTCTCTCTTTCACGCTGATCTTGGATTCTGTC 600
 Qy 601 CGAAGGAAAGAAAGATCCCTTCTGCTGCTGCGAGGTTGAGTTGACATTCATGCTCTT 660
 Db 601 CGAAGGAAAGAAAGATCCCTTCTGCTGCTGCGAGGTTGAGTTGACATTCATGCTCTT 660
 Qy 661 GTGTAATTCAGATGAGGCGCTATTTGAGGAGTGTGATCTTGGGTCTGTGACC 720
 Db 661 GTGTAATTCAGATGAGGCGCTATTTGAGGAGTGTGATCTTGGGTCTGTGACC 720
 Qy 721 ATCCGACTTCTTCTTGGGAAAGGCTGCTAA 753
 Db 721 ATCCGACTTCTTCTTGGGAAAGGCTGCTAA 753

RESULT 13

LOCUS BD014992 753 bp DNA linear PAT 27-AUG-2002
 DEFINITION BD014992
 ACCESSION BD014992
 VERSION BD014992.1 GI:2255799
 KEYWORDS
 SOURCE
 unidentified

REFERENCE	ORGANISM	unclassified.
AUTHORS		
TITLE		
JOURNAL	1 (bases 1 to 753)	
COMMENT	Kennernuknehto,N., Sahn,H., Eggering,L. and Pfefferle,W. Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof Patent: JP 2001169788-A 2 26-JUN-2001; DECUSSA HOELS AG,FORSCHUNGSZENTRUM JUELICH GMBH OS Corynebacterium glutamicum ATCC14752 PN JP 2001169788-A/2	
	PD 26-JUN-2001	
	PF 24-OCT-2000 JP 2000324315	
	PR 27-OCT-1999 DE 19951708.8	
	PI NICOLE KENNERNUKNEHTO,HERMANN SAHN,LOTHAR EGGERING,WALTER PFEFFERLE	
	PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/06//	
	PC C12N15/09,C12R1:15),(C12N1/21,C12R1:15),(C12P13/06,C12R1:15),,	
	PC C12N15/00,	
	PC C12N15/00,C12R1:15)	
	CC key	
	CC brmp	
	FE Location/Qualifiers	
	FT CDS (1)..(753).	
FEATURES	1..753	
source	Location/Qualifiers	
ORIGIN	1..753	
	/organism="unclassified"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:32644"	
Query Match	70.0%; Score 753; DB 6; Length 753;	
Best Local Similarity	100.0%; Pred. No. 5.6e-169; Indels 0; Gaps 0;	
Matches	753; Conservative 0; Mismatches 0;	
QY	1 GTGCAAAAACGCAAGATTCATTCAAGCCCTGGAGTGTGCCATCCAGGACCCCTG 60	
DB	1 GTGCAAAAACGCAAGATTCATTCAAGCCCTGGAGTGTGCCATCCAGGACCCCTG 60	
QY	61 GAACAGATGATAAAGATTATGGGGCTACGAAATCGGCGAAGTCTAAAACCTCCCTT 120	
DB	61 GAACAGATGATAAAGATTATGGGGCTACGAAATCGGCGAAGTCTAAAACCTCCCTT 120	
QY	121 GCTGAGATTGGGATGATGACCCGATTGGTATTGGGTTGATCTCTGGTTATTCATAC 180	
DB	121 GCTGAGATTGGGATGATGACCCGATTGGTATTGGGTTGATCTCTGGTTATTCATAC 180	
QY	181 GGCTACGAATGTTGGGACAGCCCACTGTTTCGGCTGATTTTGGCGGGCTCCACCGAA 240	
DB	181 GGCTACGAATGTTGGGACAGCCCACTGTTTCGGCTGATTTTGGCGGGCTCCACCGAA 240	
QY	241 ATGCTGATATGCCCTCGTTTGGGCGAGCGCCCTGGGGCGCATCGCGCTCACACA 300	
DB	241 ATGCTGATATGCCCTCGTTTGGGCGAGCGCCCTGGGGCGCATCGCGCTCACACA 300	
QY	301 TTGCTGTGTAACCTTCGCGACGATATCTATGGTTTCAATCCCGTGTGATGTGTCAA 360	
DB	301 TTGCTGTGTAACCTTCGCGACGATATCTATGGTTTCAATCCCGTGTGATGTGTCAA 360	
QY	361 AACCCATTGCCCGTTTCTATTGCGTTTTCGGCTATGACGAAGCTACAGCACTACT 420	
DB	361 AACCCATTGCCCGTTTCTATTGCGTTTTCGGCTATGACGAAGCTACAGCACTACT 420	
QY	421 GCGGCGACGCGCGAGGCTGTGTGCGGTGCGCATTTATCTCAATGCAATAGCGTTTAC 480	
DB	421 GCGGCGACGCGCGAGGCTGTGTGCGGTGCGCATTTATCTCAATGCAATAGCGTTTAC 480	
QY	481 TCTTACTGGTATTTGGCGGATCTCACCGAGTGTGCGCATGCAAGATTTGATCTTTTGA 540	
DB	481 TCTTACTGGTATTTGGCGGATCTCACCGAGTGTGCGCATGCAAGATTTGATCTTTTGA 540	
QY	541 ATTAAAGGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGAATTGATTCCTGC 600	
DB	541 ATTAAAGGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGAATTGATTCCTGC 600	

Accession	Gene	Protein	Length
QY 601	CGAAGCAAAAAGCGAGATCCCTTCCTCTGCTGCTGCGAGATTGAACTTACCATTTGGCTTT	660	
QY 601	CGAAGCAAAAAGCGAGATCCCTTCCTCTGCTGCTGCGAGATTGAACTTACCATTTGGCTTT	660	
Db 601	CGAAGCAAAAAGCGAGATCCCTTCCTCTGCTGCTGCGAGATTGAACTTACCATTTGGCTTT	660	
QY 661	GTGCTATTTCAGAGTCAGGCGCCATTATTTGCGGCGCTGCTGATCTCTTGGGCTGTGAC	720	
QY 661	GTGCTATTTCAGAGTCAGGCGCCATTATTTGCGGCGCTGCTGATCTCTTGGGCTGTGAC	720	
Db 661	GTGCTATTTCAGAGTCAGGCGCCATTATTTGCGGCGCTGCTGATCTCTTGGGCTGTGAC	720	
QY 721	ATCCGTAATCTTCTTGGGAAAGGCGTAA	753	
QY 721	ATCCGTAATCTTCTTGGGAAAGGCGTAA	753	
Db 721	ATCCGTAATCTTCTTGGGAAAGGCGTAA	753	

RESULT	14		
BD162490			
LOCUS	BD162490	753 bp	DNA
DEFINITION	Novel polynucleotide.		linear
ACCESSION	BD162490		
VERSION	BD162490.1 GI:27866248		
KEYWORDS	JP 2002191370-A/289.		
SOURCE	unidentified		
ORGANISM	unidentified.		
	unclassified.		

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 753)	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ohiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.	Novel polynucleotide	Patent: JP 2002191370-A 289 09-JUL-2002;

COMMENT
KYOWA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
BN JP 2002191370-2/289

```

EN      09-JUL-2002
PD      09-JUL-2002
PF      15-DEC-2000 JP 2000405096
PI      SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
PI      KEIKO OCHIAI,
PI      HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
PI      OZAKI
PC      C12N1/15, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
PC      C12N1/15, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC      04, C12P13/08,
PC      C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC      G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1/15),
PC      (C12N1/21, C12R1/13), (C12N1/21, C12R1/01), (C12P13/08, C12R1/15),
PC      C12N5/00, C12N15/00
PC      C12N5/00
CC      Novel polynucleotide
FH      Key
FH      Location/Qualifiers
FT      source
FT      Location/Qualifiers
FT      /organism='Corynebacterium glutamicum'.
FEATURES
Source
1..753
1..753 chr=Hwidiect1.f1041

```

ORIGIN	70.0%; Score 753; DB 6; Length 753;
Query Match	Best Local Similarity 100.0%; Pred. No. 5,6e-189; Mismatches 0; Indels 0; Gaps 0
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 GTGCAAAAACGCAAGAAATTCATTCACAGCCCTGGAGGAGTGTGGCCATCCCAAGGCAAGCCCTG 60
Db	1 GTGCAAAAACGCAAGAAATTCATTCACAGCCCTGGAGGAGTGTGGCCATCCCAAGGCAAGCCCTG 60
QY	61 GAACAGATGATAAAGTTATCGGCGCTACGAAATCGGCGCAAGTCTAAAAACCTCCCTT 120
Db	61 GAACAGATGATAAAGTTATCGGCGCTACGAAATCGGCGCAAGTCTAAAAACCTCCCTT 120
QY	121 GGTGCAAGTTTGGGATGTAACCCGATGTGATTTGGCTTGGTCTCTGTGTTTCAATAC 180
Db	121 GGTGCAAGTTTGGGATGTAACCCGATGTGATTTGGCTTGGTCTCTGTGTTTCAATAC 180

```

QY 181 GGCTACGAATGCTGGGCAAGCCCACTGTTTCCGGGCTGATTTTCGGGGGCTCCACCGAA 240
Db 181 GGCTACGAATGCTGGGCAAGCCCACTGTTTCCGGGCTGATTTTCGGGGGCTCCACCGAA 240
QY 241 ATGCTGTCATCGCCCTGCTGTGGGCGCAGCGCCCTGGGCGCATCGCCTCACACA 300
Db 241 ATGCTGTCATCGCCCTGCTGTGGGCGCAGCGCCCTGGGCGCATCGCCTCACACA 300
QY 301 TTGCTGTGAATCTCCGCGCAGATTCATGAGCTTTTCATTCCTGCTGATGTGCAAA 360
Db 301 TTGCTGTGAATCTCCGCGCAGATTCATGAGCTTTTCATTCCTGCTGATGTGCAAA 360
QY 361 AACCCATTTGCCCGTTTCTATTCGGTTTCGGGCTTATGAGAGGCTTACGAGTCACT 420
Db 361 AACCCATTTGCCCGTTTCTATTCGGTTTCGGGCTTATGAGAGGCTTACGAGTCACT 420
QY 421 GCGGCGAGGCGCGAGGCTGGTGGCGGTGCGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
Db 421 GCGGCGAGGCGCGAGGCTGGTGGCGGTGCGGCACTTATCTCAATGCAAAATAGCGTTTCA 480
QY 481 TCCGACTGGGTATTTGGGGGTCTCAGCGAGTGGGAGATCGGAGACTTATCTTTGAA 540
Db 481 TCCGACTGGGTATTTGGGGGTCTCAGCGAGTGGGAGATCGGAGACTTATCTTTGAA 540
QY 541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 600
Db 541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 600
QY 601 CGAAGAAAAGCAATCCCTTCTCTGCTGCTGCGAGTTGAGTTGACATTCCTT 660
Db 601 CGAAGAAAAGCAATCCCTTCTCTGCTGCTGCGAGTTGAGTTGACATTCCTT 660
QY 661 GTGGTAATTCAGGTCAGGCGCTATTTTGGGGGCTGATCTTCTTGGGTCTGTGACC 720
Db 661 GTGGTAATTCAGGTCAGGCGCTATTTTGGGGGCTGATCTTCTTGGGTCTGTGACC 720
QY 721 ATCCGTAATCTTCTTCTTGGGAAAAGGCTGCTAAA 753
Db 721 ATCCGTAATCTTCTTCTTGGGAAAAGGCTGCTAAA 753

```

RESULT 15

AR391955

LOCUS AR391955 324 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 4 from patent US 6613545.

ACCESSION AR391955

VERSION AR391955.1 GI:40115726

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 324)

AUTHORS Kemerzinech,N., Salm,H., Eggeling,L. and Pfeifferle,W.

TITLE Nucleotide sequences coding for the export of branched chain amino

JOURNAL acids, process for the isolation thereof and use thereof

FEATURES Location/Qualifiers

1..324

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 30.1%; Score 324; DB 6; Length 324;

Best Local Similarity 100.0%; Pred. No. 5.5e-75; Mismatches 0; Indels 0; Gaps 0;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACAACTGATTTCTCTGTAATCTCTGTTGTGCGAGTATGCAATGATTTT 812

Db 1 ATGACAACTGATTTCTCTGTAATCTCTGTTGTGCGAGTATGCAATGATTTT 60

QY 813 GGGCTCCGGGCGGTTCCGTTTATCTTAAAGCCCTAGCGAATGCAATTTGTGGC 872

```

Db 61 GCGCTCCGGGCGGTTCCGTTTATCTTAAAGCCCTAGCGAATGCAATTTGTGGC 120
QY 873 AAAATGCCATGTGATGTCGACGAGGATCTTGCCATTTTGAACGATCAAGTTTGC 932
Db 121 AAAATGCCATGTGATGTCGACGAGGATCTTGCCATTTTGAACGATCAAGTTTGC 180
QY 933 AGCAATGCCATGTGATGTCGACGAGGATCTTGCCATTTTGAACGATCAAGTTTGC 992
Db 181 AGCAATGCCATGTGATGTCGACGAGGATCTTGCCATTTTGAACGATCAAGTTTGC 240
QY 993 GTGGCGCATCTTCTTGGCGGTGACGCGCACTTGTGAGCGTTGGCGTGGCAACATCGTT 1052
Db 241 GTGGCGCATCTTCTTGGCGGTGACGCGCACTTGTGAGCGTTGGCGTGGCAACATCGTT 300
QY 1053 TTGTTGCACTGTTGAATCTTTTC 1076
Db 301 TTGTTGCACTGTTGAATCTTTTC 324

```

Search completed: April 19, 2004, 18:29:53
Job time: 4563.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 501.089 Seconds

(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176

Perfect score: 1076

Sequence: 1 gtgcacaaacgcagagat.....ctgacgctgacatcttc 1076

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2003bs:*
8: geneseqn2003cs:*
9: geneseqn2004s:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	100.0	1271	4	AAH21109 C. glutam
2	1076	100.0	1271	4	AAH21112 C. glutam
3	1076	100.0	349980	5	AAH64966 C. glutam
4	753	70.0	753	4	AAH21110 C. glutam
5	753	70.0	753	5	AAH65254 C. glutam
6	553	51.4	564	7	ACA01969 C. glutam
7	324	30.1	324	4	AAH21111 C. glutam
8	324	30.1	324	5	AAH65255 C. glutam
9	324	30.1	324	7	ACA01970 C. glutam
10	273.6	25.4	397	4	AAH61693 C. glutam
11	273.6	25.4	398	4	AAH61695 C. glutam
12	127.4	11.8	711	5	AAH68373 C. glutam
13	127.4	11.8	711	5	ACA01294 C. glutam
14	127.4	11.8	309400	5	AAH68534 C. glutam
15	106	9.9	177	7	ACA01968 C. glutam
16	85	7.9	715	4	AAH61688 C. glutam
17	48	4.5	10732	3	AAA10594 Gene enco
18	47	4.4	2000	7	ADA71938 Rice gene
19	40.8	3.8	7645	9	ADBS9610 Human gen
20	40.2	3.7	4947	2	AAQ86927 Tspst17 po
21	39.8	3.7	768	3	ADC91639 E. faeciu
22	39.2	3.6	6741	3	AAA10595 Gene enco
23	37.6	3.5	759	7	ACF69125 Photornab

C	24	37.6	3.5	5059	2	AAH64332	AAH64332 Stealth v
	25	37.6	3.5	110000	7	ACF67367_20	Continuation (21) o
	26	37.6	3.5	243072	7	ACF65382	Abt65382 Photornab
	27	37.2	3.5	1024	6	ABX6366	Abx6366 Helicobac
	28	36.8	3.4	1230	7	ACA42205	ACA42205 Prokaryot
	29	36.4	3.4	751	2	AAV58489	AAV58489 3' fragme
	30	36.4	3.4	751	2	AAV61146	AAV61146 3' CDNA s
	31	36.4	3.4	751	3	AAA06252	AAA06252 Human imm
	32	36.4	3.4	751	3	ABH71157	ABH71157 Human pro
	33	36.4	3.4	751	4	AAH93368	AAH93368 Human pro
	34	36.4	3.4	751	4	AAH93460	AAH93460 Human pro
	35	36.4	3.4	751	4	AAH02433	AAH02433 Prostate
	36	36.4	3.4	751	4	AAH94682	AAH94682 Human pro
	37	36.4	3.4	751	5	ACA59269	ACA59269 Prostate
	38	36.4	3.4	751	5	AAH00011	AAH00011 Human pro
	39	36.4	3.4	751	6	ABH94832	ABH94832 Human N1-
	40	36.4	3.4	751	6	ABH58541	ABH58541 Prostate
	41	36.4	3.4	751	7	ACC94996	ACC94996 Prostate
	42	36.4	3.4	751	9	ADB13462	ADB13462 Human pro
	43	35.6	3.3	2000	7	ADA71938	ADA71938 Rice gene
	44	35.6	3.3	4590	5	AAH24065	AAH24065 Yeast AOD
	45	35.6	3.3	135638	7	ABX34289	ABX34289 S. atrocol

ALIGNMENTS

RESULT 1
ID AAH21109
ID AAH21109 standard; DNA; 1271 BP.
AC AAH21109;
DT 05-SEP-2001 (first entry)
XX
DE C. glutamicum DNA encoding brnF and brnE.
KW L-amino acid production; brnF; brnE; branched-chain amino acid;
KW Corynebacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.
XX OS Corynebacterium glutamicum.
XX
XX Key Location/Qualifiers
FT CDS 101..856
FT FT /*tag= a
FT FT /product= "brnF"
FT CDS 853..1179
FT FT /*tag= b
FT FT /product= "brnE"
XX
XX EPI096010-A1.
XX
XX 02-MAY-2001.
XX
XX 11-OCT-2000; 2000EP-00122057.
XX
XX 27-OCT-1999; 99DE-01051708.
XX
XX (DEGS) DEGUSSA AG.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Kennernrecht N, Eggeling L, Sahm H, Pfefflerle W;
XX WPI, 2001-391595/42.
XX P-PSDB; AAB86247; AAB86248.
XX
XX New export genes from corynebacterium, useful for increasing
XX fermentative production of branched-chain amino acids.
XX Claim 4 (i); Page 13; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (i) containing

CC at least one sequence that (1) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (1i) encodes a polypeptide at least 70% identical with (3)
CC or (5); (1ii) is the complement of (1) or (1i); or (1v) contains at least
CC 15 consecutive bases from (1)-(1ii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 344 bp (4); (b)
CC coryneform microorganisms, especially *Corynebacterium*, transformed with
CC one or more (1), where these are replicative DNA; (c) production of
CC branched-chain l-a as by fermentation of coryneform bacteria in which the
CC bnfB and/or bnfF genes (or equivalent sequences) are amplified,
CC bnfB and/or overexpressed; and (d) method for isolating the bnfB and/or
CC bnfF genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC *Corynebacterium glutamicum* ATCC 14522 bnfF and bnfB proteins described in
CC the method of the invention

SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1076;	DB 4;	Length 1271;
-------------	---------	-------------	-------	--------------

Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy	GTGCAAAAACGCAAGGATTCATTCAACCTCGAAGGAGTGCCATCCAAAGCAGCCCTG	60
Db	101 GTGCAAAAACGCAAGGATTCATTCAACCTCGAAGGAGTGCCATCCAAAGCAGCCCTG	160
Qy	61 GAACCAAGATGATTAAGGTTATCGGCGCTACGAATTCGGCAGAGGTCTAAAACTCCCTT	120
Db	161 GAACCAAGATGATTAAGGTTATCGGCGCTACGAATTCGGCAGAGGTCTAAAACTCCCTT	220
Qy	121 GCTGCAGGTTTGGGCAATGTAACCGAATTTGGTATTTGGCTCTTGCTATTCATTAAC	180
Db	221 GCTGCAGGTTTGGGCAATGTAACCGAATTTGGTATTTGGCTCTTGCTATTCATTAAC	280
Qy	181 GGGTACGAATGTTGGGCAAGCCCACTGTTTCCGCTCATTTTCGCGGCTCCACCGAA	240
Db	281 GGGTACGAATGTTGGGCAAGCCCACTGTTTCCGCTCATTTTCGCGGCTCCACCGAA	340
Qy	241 ATGTGTGTCATCGCCCTCGTTGTGTGGCGCAGCGCCCTCGGCGCATATCGCCTACACA	300
Db	341 ATGTGTGTCATCGCCCTCGTTGTGTGGCGCAGCGCCCTCGGCGCATATCGCCTACACA	400
Qy	301 TTGTGTGTAACCTTCGCAACGATTCATATGAGTTTATTCCTCCGCTCATATGTGTCAA	360
Db	401 TTGTGTGTAACCTTCGCAACGATTCATATGAGTTTATTCCTCCGCTCATATGTGTCAA	460
Qy	361 AACCCCATTCGCCGTTTCTATTGCGTTTTCGCGCTTATTCAGACGATTCAGCAGTCACT	420
Db	461 AACCCCATTCGCCGTTTCTATTGCGTTTTCGCGCTTATTCAGACGATTCAGCAGTCACT	520
Qy	421 GCGGCGAGCGCCGCAAGGCTGTGCGGCGTGGGCACTTATCTCAATGCAAAATGAGCTTCAAC	480
Db	521 GCGGCGAGCGCCGCAAGGCTGTGCGGCGTGGGCACTTATCTCAATGCAAAATGAGCTTCAAC	580
Qy	481 TCTCTACTGAGTATTCGCGGCTCTCAACGAGAGTGAGCGATTCGCAAGTTGATTCCTTTGAA	540
Db	581 TCTCTACTGAGTATTCGCGGCTCTCAACGAGAGTGAGCGATTCGCAAGTTGATTCCTTTGAA	640
Qy	541 ATTAAAGGCTTCGAGTTTCGCCCTTTTGTCTCTCTTTGTCAACGTCACTTTTGGATTTCTGCG	600
Db	641 ATTAAAGGCTTCGAGTTTCGCCCTTTTGTCTCTCTTTGTCAACGTCACTTTTGGATTTCTGCG	700
Qy	501 CGAAACGAAAGACAGATTCCTTCTCTGCTGCTCGAGGTTTACGTTTACACATTTGCTCTT	660
Db	701 CGAAACGAAAGACAGATTCCTTCTCTGCTGCTCGAGGTTTACGTTTACACATTTGCTCTT	760
Qy	661 GTGTGTAATTCAGGTCAGGCTCCTATTGTGGCGGCTGTGATCTTTTGGGTCTGTGAAC	720
Db	761 GTGTGTAATTCAGGTCAGGCTCCTATTGTGGCGGCTGTGATCTTTTGGGTCTGTGAAC	820

Oy	721	ATCCGGACTCTCTCTGGGAAAAGGCTGTAATAAGCAATGATTTCTCTGTAATCTCC	780
Db	821	ATCCGGACTCTCTCTGGGAAAAGGCTGTAATAAGCAATGATTTCTCTGTAATCTCC	880
Oy	781	TTGTTGTCGAGATATGTCAGCTCAATTACTTTTGCGCTCCG99CG9TTCGGTTCTTAATCC	840
Db	881	TTGTTGTCGAGATATGTCAGCTCAATTACTTTTGCGCTCCG99CG9TTCGGTTCTTAATCC	940
Oy	841	TTAAGCCCTTAACGGAATCAACAATTTGTTGGGCAAAAATGCGATGTGATGCGACAGAGAA	900
Db	941	TTAAGCCCTTAACGGAATCAACAATTTGTTGGGCAAAAATGCGATGTGATGCGACAGAGAA	1000
Oy	901	TCCTTGGCAATTTTGACCGCATCAACGTTTGGCAGACGAATGCGATGATCTGAAGACTCTAA	960
Db	1001	TCCTTGGCAATTTTGACCGCATCAACGTTTGGCAGACGAATGCGATGATCTGAAGACTCTAA	1060
Oy	961	CCCTTGGATTCATTTGACCGGTTGCGAATTAACAGTGGTGGGCAATCTTCTTGGCGGTGAGCGCA	1020
Db	1061	CCCTTGGATTCATTTGACCGGTTGCGAATTAACAGTGGTGGGCAATCTTCTTGGCGGTGAGCGCA	1120
Oy	1021	CCCTTGTGAGCGTTGAGCGCTGCGACCAATCGTTTGTGTTGACTGTGAATCTTTTC	1076
Db	1121	CCCTTGTGAGCGTTGAGCGCTGCGACCAATCGTTTGTGTTGACTGTGAATCTTTTC	1176
RESULT 2			
AAH21112	AAH21112 standard; DNA, 1271 BP.		
XX	AAH21112;		
AC			
XX			
XX	05-SEP-2001 (first entry)		
DT			
DE	C. glutamicum DNA encoding brnF and brnE.		
XX			
XX	L-amino acid production; brnF, brnE; branched-chain amino acid;		
KV	coryneform bacterium; leucine; isoleucine; valine; medicine;		
KW	animal nutrition; ds.		
XX			
OS	Corynebacterium glutamicum.		
XX			
XX	Key	Location/Qualifiers	
FH	CDS	101..856	
FT		/*tag= a	
FT		/product= "brnF"	
FT	CDS	853..1179	
FT		/*tag= b	
FT		/product= "brnE"	
XX			
XX	EP1096010-A1.		
PN			
XX			
PD	02-MAY-2001.		
XX			
PF	11-OCT-2000; 2000EP-00122057.		
XX			
PR	27-OCT-1999; 99DE-01051708.		
XX			
XX	(DEGS) DEGUSSA AG.		
PA	(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.		
XX			
PI	Kennernrecht N, Eggeling L, Salm H, Pfeifferle W,		
XX			
DR	WPI; 2001-391595/42.		
XX			
XX	New export genes from coryneform bacteria, useful for increasing		
PT	fermentative production of branched-chain amino acids.		
XX			
XX	Claim 4 (i); Page 17-18; 23pp; German.		
CC	This invention describes a novel isolated polynucleotide (I) containing		
CC	at least one sequence that (1) is 70% identical with a sequence that		
CC	encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)		

This invention describes a novel isolated polynucleotide (I) containing at least one sequence that (1) is 70% identical with a sequence that encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)

CC in electronic format directly from the European Patent Office

xx Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;

Query Match 100.0%; Score 1076; DB 5; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCATCCAGAGCCCTG 60
DB 276829 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCATCCAGAGCCCTG 276888
QY 61 GAACCAAGATGATTAAGTTATGCGCTACGAAAATCGCGCAAGCTTAAAAACCTCCCTT 120
DB 276889 GAACCAAGATGATTAAGTTATGCGCTACGAAAATCGCGCAAGCTTAAAAACCTCCCTT 276948
QY 121 GCTGCAAGTTGGGCGATGACCCGATGATGATGCTTGGCTCTCTTGGTTATTCATAC 180
DB 276949 GCTGCAAGTTGGGCGATGATGATGATGCTTGGCTCTCTTGGTTATTCATAC 277008
QY 181 GACTCAAGATGGTGGGCAAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 240
DB 277009 GACTCAAGATGGTGGGCAAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 277068
QY 241 ATGCTGATCATGCGCTTGGTGGGCGCAAGCCCTCGGGCGCCATCGGCTCAACCA 300
DB 277069 ATGCTGATCATGCGCTTGGTGGGCGCAAGCCCTCGGGCGCCATCGGCTCAACCA 277128
QY 301 TTGCTGATGATTCGCGCAAGTATTCATGATGCTTTTCAATCCCGCTGATGATGATCA 360
DB 277129 TTGCTGATGATTCGCGCAAGTATTCATGATGCTTTTCAATCCCGCTGATGATGATCA 277188
QY 361 AACCCATTGCGCGTTCATTCGCTTTCGCGCTTATCGAAGAGCTTACGAGTCACT 420
DB 277189 AACCCATTGCGCGTTCATTCGCTTTCGCGCTTATCGAAGAGCTTACGAGTCACT 277248
QY 421 GCGGCAAGCCCGGAGCTGCTGCGCGCTGCGGCACTTATCAATGCAATGCGCTTAC 480
DB 277249 GCGGCAAGCCCGGAGCTGCTGCGCGCTGCGGCACTTATCAATGCAATGCGCTTAC 277308
QY 481 TCCCTAGGATTCGCGGCTGCTCAACGAGTGGCGATGCAAGATGATTCCTTTGAA 540
DB 277309 TCCCTAGGATTCGCGGCTGCTCAACGAGTGGCGATGCAAGATGATTCCTTTGAA 277368
QY 541 ATTAAGGCGCTTGAAGTTCCTTCTCTCTTCTTCTGCAAGCTTACCTTTCCTGCG 600
DB 277369 ATTAAGGCGCTTGAAGTTCCTTCTCTCTTCTTCTGCAAGCTTACCTTTCCTGCG 277428
QY 601 CGAAGCAAAAACGATCCCTTCTGCTGCTGCGAGGTTGAGCTTACCATTCCTCTT 660
DB 277429 CGAAGCAAAAACGATCCCTTCTGCTGCTGCGAGGTTGAGCTTACCATTCCTCTT 277488
QY 661 GTGCTAATTCAGGTCAGGCTTATTTGGGCGCTGCTGATTTTGGGCTGTTGACC 720
DB 277489 GTGCTAATTCAGGTCAGGCTTATTTGGGCGCTGCTGATTTTGGGCTGTTGACC 277548
QY 721 ATCCGCTATCTTCTTGGGAAAGGCTGCTAATGACATGATTTCTCCGATATTCGC 780
DB 277549 ATCCGCTATCTTCTTGGGAAAGGCTGCTAATGACATGATTTCTCCGATATTCGC 277608
QY 781 TTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 277609 TTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277668
QY 841 TTAAGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 277669 TTAAGCCCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277728
QY 901 TCCCTGCAATTTTGAACCGCATCAAGCTTTCGAGCAATGCGATGATGATGATGATG 960
DB 277729 TCCCTGCAATTTTGAACCGCATCAAGCTTTCGAGCAATGCGATGATGATGATGATG 277788
QY 961 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020

DB 277789 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277848
QY 1021 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1076
DB 277849 CCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277904

RESULT 4

AAH21110
ID AAH21110 standard; DNA; 753 BP.

XX AAH21110;

XX 05-SEP-2001 (first entry)

DE C. glutamicum bnf DNA.

XX L-amino acid production; bnf; bnf; branched-chain amino acid;

KM coryneform bacterium; leucine; isoleucine; valine; medicine;

XX animal nutrition; ds.

OS Corynebacterium glutamicum.

XX EPI096010-A1.

XX 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

XX 27-OCT-1999; 99DE-01051708.

PR (DEGS) DEGUSA AG.

PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Kennerknecht N, Eggeling L, Sahn H, Pfefferle W;

DR WPI: 2001-381595/42.

XX P-PSDB; AAB86247.

PT New export genes from coryneform bacteria, useful for increasing

XX fermentative production of branched-chain amino acids.

PS Claim 5; Page 14-15; 23pp; German.

XX This invention describes a novel isolated polynucleotide (I) containing

CC at least one sequence that (i) is 70% identical with a sequence that

CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)

CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)

CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least

CC 15 consecutive bases from (i)-(iii). The invention also describes (a)

CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)

CC coryneform microorganisms, especially Corynebacterium, transformed with

CC one or more (I), where these are replicative DNA; (c) production of the

CC branched-chain L-aa by fermentation of coryneform bacteria in which the

CC bnf and/or bnf genes (or equivalent sequences) are amplified,

CC especially overexpressed; and (d) method for isolating the bnf and/or

CC bnf genes. (i) is used for transformation of coryneform bacteria being

CC used for fermentative production of branched-chain amino acids,

CC specifically leucine, isoleucine and valine, which are useful in medicine

CC and animal nutrition. (i) can also be used as source of primers and

CC probes for isolation of related sequences. Transformation with (i)

CC increases yield of branched-chain amino acids. This sequence encodes the

CC Corynebacterium glutamicum ATCC 14752 bnf protein described in the

CC method of the invention

XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

QY Query Match 70.0%; Score 753; DB 4; Length 753;

Best Local Similarity 100.0%; Pred. No. 8; 1e-227;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCATCCAGAGCCCTG 60

```

Db      1 GTGCAAAAAGCAGAGATTCATTCAAGCCTGAGGTGTCGATCCAGGACGCCCTG 60
Qy      61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGCTTAAACCTCCCTT 120
Db      61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGCTTAAACCTCCCTT 120
Qy      121 GCTCAGGTTGGGCAATGATCCGATGTTATGCTGTTGCTCTTGGTTATCAATAC 180
Db      121 GCTCAGGTTGGGCAATGATCCGATGTTATGCTGTTGCTCTTGGTTATCAATAC 180
Qy      181 GGCTACGAATGATGAGCGACGCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACGAA 240
Db      181 GGCTACGAATGATGAGCGACGCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACGAA 240
Qy      241 ATGCTGATTCGCGCTGTTGTTGGGCGCAGCGCCCTGGGCGGCATCGGCTCACACA 300
Db      241 ATGCTGATTCGCGCTGTTGTTGGGCGCAGCGCCCTGGGCGGCATCGGCTCACACA 300
Qy      301 TTGCTGATGATTCGCGCAATGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 360
Db      301 TTGCTGATGATTCGCGCAATGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 360
Qy      361 AACCCATGCGCGCTGTTGATTCGCGTTTGGCGCTTATCGAAGGCTTACGAGTCACT 420
Db      361 AACCCATGCGCGCTGTTGATTCGCGTTTGGCGCTTATCGAAGGCTTACGAGTCACT 420
Qy      421 GCGGCGAGCGCGAGGCTGCTGCGCGGCGGCACTTATTCGATTCGATTCGATTCG 480
Db      421 GCGGCGAGCGCGAGGCTGCTGCGCGGCGGCACTTATTCGATTCGATTCGATTCG 480
Qy      481 TCCTACTGGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 540
Db      481 TCCTACTGGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 540
Qy      541 ATTAAAGGCGCGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 600
Db      541 ATTAAAGGCGCGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 600
Qy      601 CGAAGCAAAAGCAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTCGCTT 660
Db      601 CGAAGCAAAAGCAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTCGCTT 660
Qy      661 GTGTAATTCAGGTCAGGCTTATTCGCGCGCTGCTGATCTTCTTGGGTTGTTACC 720
Db      661 GTGTAATTCAGGTCAGGCTTATTCGCGCGCTGCTGATCTTCTTGGGTTGTTACC 720
Qy      721 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 753
Db      721 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 753

```

RESULT 5
AAH65254 standard; DNA; 753 BP.

AAH65254;
26-SEP-2001 (first entry)
C glutamicum coding sequence fragment SEQ ID NO: 289.
Corynebacterium; amino acid synthesis; vitamin; saccharide;
organic acid synthesis; de.
Corynebacterium glutamicum.
EPI108790-A2.
20-JUN-2001.
18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99UP-00377484.
PR 07-APR-2000; 2000JP-00159163.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
PI Tateishi N, Senoh A, Ikeda M, Ozaki A,
XX
XX MPI, 2001-376931/40.
DR P-PSDB; AAG30035.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 289; 246bp + Sequence Listing; English.

XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacteria Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacteria, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium bacteria, and identifying a homologue of a gene derived from
CC Corynebacterium bacteria. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

XX
SQ Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

XX
Query Match 70.0%; Score 753; DB 5; Length 753;

XX
Best Local Similarity 100.0%; Pred. No. 8,1e-227;

XX
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 GTGCAAAAAGCAGAGATTCATTCAAGCCTGAGGTGTCGATCCAGGACGCCCTG 60
Db      1 GTGCAAAAAGCAGAGATTCATTCAAGCCTGAGGTGTCGATCCAGGACGCCCTG 60
Qy      61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGCTTAAACCTCCCTT 120
Db      61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGCTTAAACCTCCCTT 120
Qy      121 GCTCAGGTTGGGCAATGATCCGATGTTATGCTGTTGCTCTTGGTTATCAATAC 180
Db      121 GCTCAGGTTGGGCAATGATCCGATGTTATGCTGTTGCTCTTGGTTATCAATAC 180
Qy      181 GGCTACGAATGATGAGCGACGCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACGAA 240
Db      181 GGCTACGAATGATGAGCGACGCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACGAA 240
Qy      241 ATGCTGATTCGCGCTGTTGTTGGGCGCAGCGCCCTGGGCGGCATCGGCTCACACA 300
Db      241 ATGCTGATTCGCGCTGTTGTTGGGCGCAGCGCCCTGGGCGGCATCGGCTCACACA 300
Qy      301 TTGCTGATGATTCGCGCAATGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 360
Db      301 TTGCTGATGATTCGCGCAATGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 360
Qy      361 AACCCATGCGCGCTGTTGATTCGCGTTTGGCGCTTATCGAAGGCTTACGAGTCACT 420
Db      361 AACCCATGCGCGCTGTTGATTCGCGTTTGGCGCTTATCGAAGGCTTACGAGTCACT 420
Qy      421 GCGGCGAGCGCGAGGCTGCTGCGCGGCGGCACTTATTCGATTCGATTCGATTCG 480
Db      421 GCGGCGAGCGCGAGGCTGCTGCGCGGCGGCACTTATTCGATTCGATTCGATTCG 480
Qy      481 TCCTACTGGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 540
Db      481 TCCTACTGGGATTCGCGCGGCTGTCACCGAGTGGCGATTCGAGATTCGATTCG 540

```


CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i) - (iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (i), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC hmb and/or hmb genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the hmb and/or
 CC hmb genes. (i) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (i) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 hmb protein described in the
 CC method of the invention

Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 30.1%; Score 324; DB 4; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.5e-91;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACACGATTTCTCCTGATTCCTGTTGTGCGAGTATGCGAGTACTTCTTT 812
 DB 1 ATGACACGATTTCTCCTGATTCCTGTTGTGCGAGTATGCGAGTACTTCTTT 60
 QY 813 GCGCTCGGCGGTCCTGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGCGGC 872
 DB 61 GCGCTCGGCGGTCCTGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGCGGC 120
 QY 873 AAAATGCGGATGATGATGCGAGCAAGATCCTTGCCATTGACCGCATCAAGTTTGGC 932
 DB 121 AAAATGCGGATGATGATGCGAGCAAGATCCTTGCCATTGACCGCATCAAGTTTGGC 180
 QY 933 AGCAATGCGATAGATCTGAAGACTTAACCTTTGCTCATTCGCTTGCGATTACAGTG 992
 DB 181 AGCAATGCGATAGATCTGAAGACTTAACCTTTGCTCATTCGCTTGCGATTACAGTG 240
 QY 993 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTTGAGGCTTGCGCGCATTCGTT 1052
 DB 241 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTTGAGGCTTGCGCGCATTCGTT 300
 QY 1053 TTTGTTGACCTGATGATCTTTTC 1076
 DB 301 TTTGTTGACCTGATGATCTTTTC 324

RESULT 8

AAH65255
 ID AAH65255 standard; DNA; 324 BP.

AAH65255;

AC 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 290.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis; ds.

XX Corynebacterium glutamicum.

OS EP1108790-A2.

PN 20-JUN-2001.

PD 18-DEC-2000; 2000EP-00127688.

PF 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280985.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

DR P-PSDB; AAG90036.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

PS Claim 8; SEQ ID NO 290; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from

CC Coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a nucleic acid described

CC in the exemplification of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the European Patent Office

Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 30.1%; Score 324; DB 5; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.5e-91;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACACGATTTCTCCTGATTCCTGTTGTGCGAGTATGCGAGTACTTCTTT 812
 DB 1 ATGACACGATTTCTCCTGATTCCTGTTGTGCGAGTATGCGAGTACTTCTTT 60
 QY 813 GCGCTCGGCGGTCCTGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGCGGC 872
 DB 61 GCGCTCGGCGGTCCTGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGCGGC 120
 QY 873 AAAATGCGGATGATGATGCGAGCAAGATCCTTGCCATTGACCGCATCAAGTTTGGC 932
 DB 121 AAAATGCGGATGATGATGCGAGCAAGATCCTTGCCATTGACCGCATCAAGTTTGGC 180
 QY 933 AGCAATGCGATAGATCTGAAGACTTAACCTTTGCTCATTCGCTTGCGATTACAGTG 992
 DB 181 AGCAATGCGATAGATCTGAAGACTTAACCTTTGCTCATTCGCTTGCGATTACAGTG 240
 QY 993 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTTGAGGCTTGCGCGCATTCGTT 1052
 DB 241 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTTGAGGCTTGCGCGCATTCGTT 300
 QY 1053 TTTGTTGACCTGATGATCTTTTC 1076
 DB 301 TTTGTTGACCTGATGATCTTTTC 324

RESULT 9

ACA01970
 ID ACA01970 standard; DNA; 324 BP.

ACA01970;

AC 04-JUN-2003 (first entry)

DE C. glutamicum derived ORF SEQ ID 1961.

XX Coryneform; nucleic acid array; fermentation; culture; ds.

OS Corynebacterium glutamicum.
XX DE10128510-A1.
XX
XX 19-DEC-2002.
XX
XX 13-JUN-2001; 2001DE-01028510.
XX
XX 13-JUN-2001; 2001DE-01028510.
XX
XX (DEGS) DEGUSA AG.
XX
XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
XX WPI; 2003-279970/28.
XX
XX New nucleic acid array useful for monitoring mRNA expression of
XX Corynebacterium glutamicum during fermentation, comprising nucleic acid
XX from Corynebacterium glutamicum.
XX
XX Claim 1; Page 647; 709pp; German.
XX
XX This invention describes a novel nucleic acid array involving
XX Corynebacterium glutamicum polynucleotides. The arrays are used to
XX analyze C. glutamicum, particularly for monitoring a fermentation process
XX to determine expression levels of C. glutamicum cellular mRNA. Such
XX monitoring particularly differentiates between expression levels of
XX different strains of C. glutamicum and allows the adjustment of different
XX culture and fermentation conditions. ACA00010-ACA02188 represent C.
XX glutamicum derived polynucleotides described in the disclosure of the
XX invention
XX
XX Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
XX
XX Query Match 30.1%; Score 324; DB 7; Length 324;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-51;
XX Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 753 ATGACACTGATTTCTCTGATATTCCTGTTGTCAGATGTCAGTCATTAATTT 812
DB 1 ATGACACTGATTTCTCTGATATTCCTGTTGTCAGATGTCAGTCATTAATTT 60
QY 813 GCGCTCCGGCGGCTTCCGTTTAACTTAAGCCCTCCGTAACAATTTGCGGC 872
DB 61 GCGCTCCGGCGGCTTCCGTTTAACTTAAGCCCTCCGTAACAATTTGCGGC 120
QY 873 AAAATGCGATGTGATGTCAGAGAGATCTTGCATTTTGAACGCATCAACGTTTGC 932
DB 121 AAAATGCGATGTGATGTCAGAGAGATCTTGCATTTTGAACGCATCAACGTTTGC 180
QY 933 AGCAATGCCATATCTGAAGACTTAACTTTGTCATTCCTTCCGATTAACAGTG 992
DB 181 AGCAATGCCATATCTGAAGACTTAACTTTGTCATTCCTTCCGATTAACAGTG 240
QY 993 GTGGCGCATTTCTTGGCGGTCAGACGACCTTTGAGCTTGGCGCTGACACATGTT 1052
DB 241 GTGGCGCATTTCTTGGCGGTCAGACGACCTTTGAGCTTGGCGCTGACACATGTT 300
QY 1053 TTTGTTGACTGTGAATCTTTTC 1076
DB 301 TTTGTTGACTGTGAATCTTTTC 324

RESULT 10
AAF61693/c
ID AAF61693 standard; DNA; 397 BP.
XX
XX AAF61693;
XX
XX 12-JUL-2001 (first entry)
XX
XX C. glutamicum lrp encoding DNA fragment SEQ ID 7.
XX

KW Fermentation; L-lysine acid production; lrp gene; lysine; isoleucine;
KW medicine; animal feed supplement; ds.
XX
XX Corynebacterium glutamicum.
XX
XX EPI090993-A1.
XX
XX 11-APR-2001.
XX
XX 29-SEP-2000; 2000EP-00121159.
XX
XX 05-OCT-1999; 99DE-01047792.
XX
XX (DEGS) DEGUSA-HUELS AG.
XX
XX Moeckel B, Pfefferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX
XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX
XX Example 3; Page 16; 22pp; German.
XX
XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention
XX
XX Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;
XX
XX Query Match 25.4%; Score 273.6; DB 4; Length 397;
XX Best Local Similarity 98.6%; Pred. No. 1.4e-75;
XX Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCAGAGCCCTG 60
DB 284 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCAGAGCCCTG 225
QY 61 GACCAAGATGATAAGGTATTCGGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 120
DB 224 GACCAAGATGATAAGGTATTCGGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 165
QY 121 GCTGCAAGTTGGGCAATGACCGGATGTGATTCGTTGCTCTGTTATTAATAC 180
DB 164 GCTGCAAGTTGGGCAATGACCGGATGTGATTCGTTGCTCTGTTATTAATAC 105
QY 181 GGTACGATGATGGGAGAGCCCACTGTTTCCGCGCTGATTTGCGGAGCTCCACGAA 240
DB 104 GGTACGATGATGGGAGAGCCCACTGTTTCCGCGCTGATTTGCGGAGCTCCACGAA 45
QY 241 ATGCTGTGATCGCCCTGCTGTTGTTGGGCGACGCGCCCTCG 280
DB 44 ATGCTGTGATCGCCCTGCTGTTGTTGGGCGACGCGCCCTCG 5

RESULT 11
AAF61695/c
ID AAF61695 standard; DNA; 778 BP.
XX
XX AAF61695;
XX
XX 12-JUL-2001 (first entry)
XX
XX C. glutamicum lrp encoding DNA fragment SEQ ID 9.
XX

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KW medicine; animal feed supplement; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EPI090993-A1.
 XX
 PD 11-APR-2001.
 XX
 PF 29-SEP-2000; 2000EP-00121159.
 XX
 PR 05-OCT-1999; 99DE-01047792.
 XX
 PA (DEGS) DEGUSA-HUELS AG.
 XX
 PI Moeckel B, Pfeifferle W, Fuehler A, Kalinowski J, Bathe B;
 XX
 DR WPI; 2001-292927/31.
 XX
 PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.
 XX
 PS Example 3; Page 17; 22pp; German.
 XX
 CC This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention
 XX
 SQ Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;
 Query Match 25.4%; Score 273.6; DB 4; Length 778;
 Best Local Similarity 98.6%; Pred. No. 1.9e-75;
 Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGAGCCCTG 60
 DB 284 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGAGCCCTG 225
 QY 61 GAACCAATGATTAAGTTATCGGCGCTAGAAATCGCGAAGTCTAAAAACCTCCCTT 120
 DB 224 GAACCAATGATTAAGTTATCGGCGCTAGAAATCGCGAAGTCTAAAAACCTCCCTT 165
 QY 121 GCTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTGTTATTCATATAC 180
 DB 164 GCTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTGTTATTCATATAC 105
 QY 181 GGCTACGAATGTGTGGGACGCCCACTGTTTCGGGCTGATTTTCGGGGCTCCACGAA 240
 DB 104 GGCTACGAATGTGTGGGACGCCCACTGTTTCGGGCTGATTTTCGGGGCTCCACGAA 45
 QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGACGCGCCCTGG 280
 DB 44 ATGCTGTCATCGCCCTGTTGTGGGCGACGCGCCCGGG 5

DE C glutamicum coding sequence fragment SEQ ID NO: 3408.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EPI08790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 XX
 PR 07-APR-2000; 2000JP-00159162.
 XX
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 XX
 DR F-PSDB; AAG93154.
 XX
 PT Novel polynucleotides derived from coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 3408; 246bp + Sequence listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 11.8%; Score 127.4; DB 5; Length 711;
 Best Local Similarity 51.5%; Pred. No. 2.7e-29;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 122 GTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTGTTATTCATATAC 181
 DB 53 CGATGCGTTGGGCGGTGATTCGCTGGGTTGGCTTTGGGTGTGATGTCAGACAG 112
 QY 182 GCTACGAATGTGTGGGACGCCCACTGTTTCGGGCTGATTTTCGGGCTCCACGAA 241
 DB 113 GTTTGCGCTGTGTGTGGAGCGCGATTTCTCTTGTGATCTATGCGGTTGATGAAAT 172
 QY 242 TCGTGTATCGCCCTGTTGTGGGCGACGCCCTTGGGCGCATCGGCTCACCAAT 301
 DB 173 TTCTGCAATCGCATGTGACCCGAGATTCGCGCTTTTCGGGCGGCTGTGGTT 232
 QY 302 TCGTGTGAATTCGCGACGATTCATGAGTTTCATTCGCGGTGATGATGATCAAA 361
 DB 233 TCAATGTGATTTCCGCAATTTTCTACGTTCTACCTTCCACGCAACGCAATCAAT 292
 QY 362 ACCCAATGCGCGTTTCTATTTGATTTTCGGCTTATCGACGAACCTTACGATCACTG 421
 DB 293 CCGGCGCGCGCGCGCTATTCACCTACGCGCTTACGAGAGCTTACGCGCATGCTGT 352
 QY 422 CGGCGAGCGCGGAGCTGTGTGGGCTG---CGGACTTATCTCATGGAATAGGTTTC 478
 DB 353 CAGCGCGCGCGCATTCGATTCAGTGTGACGCGGCTCTTACGTTCAATTTTGTGCC 412

RESULT 12
 AAH68373
 ID AAH68373 standard; DNA; 711 BP.
 XX
 AC AAH68373;
 XX
 DT 26-SEP-2001 (first entry)

QY 479 ACTCTACTGGGATTTGCGCGCTCTCACCGAGTGGCGATCGAGATGATTCCTTTG 538
 Db 413 AAGCTCTGTGGGTTATCCAGGAATTAATTGGCGCTTGTGTGCAAGTGTGCGCATG 472
 QY 539 AAATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTACAGCTGATTTGATTCCT 598
 Db 473 ATCTAAAGGATGATTTTGTGCTCGACCGCGCTGTTGTGTCTGCGCGTGGAGGAT 532
 QY 599 GCCGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTC 658
 Db 533 TCAAAATATACAGAGATTAATTCGTCGATTAATTCGCGATGATGAGCTTGGTTCCG 592
 QY 659 TTGTGTAATTCGAGAGCGCCATTTGCGGCGCTGCTGATCTTCTGTGGTGTGTA 718
 Db 593 GTTTGTGGCGCCGACGAGATGCTGTATGCTTTGACACGATCTTTTATCTTC 652
 QY 719 CCATCCGCTACTTCTTC 735
 Db 653 TTCTCCGCTCCGCTTC 669

RESULT 13

ACA01294
 ID ACA01294 standard; DNA; 711 BP.

ACA01294;

DT 03-JUN-2003 (first entry)

DE C. glutamicum derived ORF SEQ ID 1285.

KM Coryneform; nucleic acid array; fermentation; culture; ds.

OS Corynebacterium glutamicum.

PN DE10128510-A1.

PD 19-DEC-2002.

PF 13-JUN-2001; 2001DE-01028510.

PR 13-JUN-2001; 2001DE-01028510.

PA (DBGS) DEGUSA AG.

PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;

DR WPI; 2003-279970/28.

PT New nucleic acid array useful for monitoring mRNA expression of
 Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.

CC Claim 1; Page 456; 709pp; German.

CC This invention describes a novel nucleic acid array involving
 Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;

XX Query Match 11.8%; Score 127.4; DB 7; Length 711;

XX Best Local Similarity 51.5%; Pred. No. 2.7e-29;

XX Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGCAAGTGTGGGCAATGACCCGATTTGGTATTCGTTGCTCTTGGTATTCAATACG 181

Db 53 CGGTGCGTTTGGAGGCTGATTCGCTGGGTTTGGCTTTGGGCTGTTATGATCCAGACG 112
 QY 182 GCTACGATGATGAGGACAGCCCACTGTTTTCGGGCTGATTTTCGGGGCTCCACCGAA 241
 Db 113 GTTTGGCTGTGTGAGAGAGCCGATTTTCTCTTGTGATCATGATCCGTTGATGAAAT 172
 QY 242 TGTGTGTAATTCGCGCTGTTGTGGGCGAGCGCCCTGGGGCCATTCGGGCTACCAAT 301
 Db 173 TTTTGGCAATGGCAATGATTCACCGAGATGATGAGCCCTTTTGGCGGCTGCTGATT 232
 QY 302 TGTGTGTAATTCGCGCAAGATTTCTATGCGTTTCAATTCGCGCTGATGTTGTA 361
 Db 233 TATGTGTAATTTCCGCCATTTTCTAGGCTCACTCCACGCGCAACGATCAAGT 292
 QY 362 ACCCATTCGCGCTTCTATTTGCTTTGCGGCTTATGAGAGGCTAAGCAATCACTG 421
 Db 293 CGGCGCGCGCGCGCTTATTCACCTTACGCGCTTACGAGATCTTACGCGCATGCTGT 352
 QY 422 CGGCAAGCGCCGACAGCTGTGCGGCTG--GCACTTATCTCAATGCAATAGCGTTTC 478
 Db 353 CAGCCGCCACCTGGCATATCAGGAGAGCGGGTGTCTTACGTTCAAAATTTGTGCG 412
 QY 479 ACTCTACTGGGATTTGCGCGCTCTCACCGAGTGGCGATCGAGATGATTCCTTTG 538
 Db 413 AAGCTCTGTGGGTTATCCAGGAATTAATTGGCGCTTGTGTGCAAGTGTGCGCATG 472
 QY 539 AAATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTACAGCTGATTTGATTCCT 598
 Db 473 ATCTAAAGGATGATTTTGTGCTCGACCGCGCTGTTGTGTCTGCGCGTGGAGGAT 532
 QY 599 GCCGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTC 658
 Db 533 TCAAAATATACAGAGATTAATTCGTCGATTAATTCGCGATGATGAGCTTGGTTCCG 592
 QY 659 TTGTGTAATTCGAGAGCGCCATTTGCGGCGCTGCTGATCTTCTGTGGTGTGTA 718
 Db 593 GTTTGTGGCGCCGACGAGATGCTGTATGCTTTTGAACGATCTTTTGAATCTTC 652
 QY 719 CCATCCGCTACTTCTTC 735
 Db 653 TTCTCCGCTCCGCTTC 669

RESULT 14

AAH68534/c
 ID AAH68534 standard; DNA; 309400 BP.

XX AAH68534;

XX 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7069.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

OS Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 90.5988 seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176
Perfect score: 1076
Sequence: 1 gtgcataaaacgaagatc.....ctgactgtgtaatcttc 1076

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 662709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1076	100.0	1271	4 US-09-471-803A-1	Sequence 1, Appl1
2	1076	100.0	1271	4 US-09-471-803A-6	Sequence 6, Appl1
3	753	70.0	753	4 US-09-471-803A-2	Sequence 2, Appl1
4	324	30.1	324	4 US-09-471-803A-4	Sequence 4, Appl1
5	58.6	5.4	756	4 US-09-469-039A-3016	Sequence 3016, Ap
6	52.4	4.9	7218	1 US-08-233-443-14	Sequence 14, Appl
7	39.8	3.7	768	4 US-09-107-532A-1266	Sequence 1266, Ap
8	39.2	3.6	1086	4 US-09-253-991A-11376	Sequence 11376, A
9	39.2	3.6	1086	4 US-09-253-991A-11418	Sequence 11418, A
10	39	3.6	399	4 US-09-621-976-8976	Sequence 8976, Ap
11	37	3.4	474	4 US-09-621-976-18033	Sequence 18033, A
12	36.8	3.4	1140	4 US-09-253-991A-3205	Sequence 3205, Ap
13	36.8	3.4	2061	4 US-09-253-991A-2761	Sequence 2761, Ap
14	36.4	3.4	751	3 US-09-020-956-12	Sequence 12, Appl
15	36.4	3.4	751	3 US-09-030-607-12	Sequence 12, Appl
16	36.4	3.4	751	3 US-09-433-313-12	Sequence 12, Appl
17	36.4	3.4	751	4 US-09-353-616A-12	Sequence 12, Appl
18	36.4	3.4	751	4 US-09-233-149A-12	Sequence 12, Appl
19	36.4	3.4	751	4 US-09-159-812-12	Sequence 12, Appl
20	36.4	3.4	751	4 US-09-636-215-12	Sequence 12, Appl
21	36.4	3.4	751	4 US-09-685-166A-12	Sequence 12, Appl
22	36.4	3.4	751	4 US-09-115-453-12	Sequence 12, Appl
23	36.4	3.4	751	4 US-09-688-489-12	Sequence 12, Appl
24	35.4	3.3	430	4 US-09-621-976-16656	Sequence 16656, A
25	35.2	3.3	801	4 US-09-328-352-2399	Sequence 2399, A
26	35.2	3.3	4403765	3 US-09-103-840A-2	Sequence 2, Appl1
27	35.2	3.3	4411529	3 US-09-103-840A-1	Sequence 1, Appl1

28	34.8	3.2	867	4 US-09-489-039A-5047	Sequence 5047, Ap
29	34.6	3.2	666	2 US-08-875-034A-1	Sequence 1, Appl
30	34.4	3.2	744	4 US-09-489-039A-1307	Sequence 1307, Appl
31	34.4	3.2	1224	4 US-09-266-965-22	Sequence 22, Appl
32	34.4	3.2	12249	4 US-09-266-965-74	Sequence 74, Appl
33	34.4	3.2	18331	4 US-09-266-965-96	Sequence 96, Appl
34	34.2	3.2	594	4 US-09-253-991A-13729	Sequence 13729, A
35	34.2	3.2	678	4 US-09-253-991A-13832	Sequence 13832, A
36	33.4	3.1	1875	3 US-08-423-869-21	Sequence 21, Appl
37	33.4	3.1	2949	3 US-08-433-522A-1	Sequence 1, Appl1
38	33.4	3.1	2849	3 US-09-135-166-1	Sequence 1, Appl1
39	33.4	3.1	2849	3 US-08-942-046-1	Sequence 1, Appl1
40	33.4	3.1	2950	3 US-08-433-522A-5	Sequence 5, Appl1
41	33.4	3.1	2950	3 US-09-135-166-5	Sequence 5, Appl1
42	33.4	3.1	2950	3 US-08-942-046-5	Sequence 5, Appl1
43	33.4	3.1	2984	3 US-08-433-522A-3	Sequence 3, Appl1
44	33.4	3.1	2984	3 US-09-135-166-3	Sequence 3, Appl1
45	33.4	3.1	2984	3 US-08-942-046-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-471-803A-1
Sequence 1, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
ISOLATION THEREOF AND USE THEREOF
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471,803A
PRIORITY FILING DATE: 1999-12-23
PRIORITY APPLICATION NUMBER: DE 199 51 708.8
PRIORITY FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnF
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnE
OTHER INFORMATION: ATCC14752
US-09-471-803A-1
Query Match 100.0%; Score 1076; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAACCGAAGATTCATTCAAGCTCGAGGTGCGCCATCCAGGACGCCCTG 60
DB 101 GTGCAAAAACCGAAGATTCATTCAAGCTCGAGGTGCGCCATCCAGGACGCCCTG 160
QY 61 GAACGAGATGATTAAGGTATCGGCGCTGCAAGATCGCAAGTCTAAACCTCCCTT 120
DB 161 GAACGAGATGATTAAGGTATCGGCGCTGCAAGATCGCAAGTCTAAACCTCCCTT 220
QY 121 GTCGAGGTTTGCGCATGTACCGGATGTGATGCTTGTCTCTGTTATTCATAC 180
DB 221 GTCGAGGTTTGCGCATGTACCGGATGTGATGCTTGTCTCTGTTATTCATAC 280
QY 181 GGTACGAATGTGGGACGCCCACTGTTTCGGGCTGATTTTGGGGGCTCCACCGAA 240

```

Db 281 GGCACAGATGATGGGAGGCCCACTGTTTCGGGGCTGATTTTCGGGGCTCCACGAA 340
Qy 241 ATGCTGATCATGCGCCCTGTTGTGGGCGACGCGCCCTGGGCGCATCGGCTCACCACA 300
Db 341 ATGCTGATCATGCGCCCTGTTGTGGGCGACGCGCCCTGGGCGCATCGGCTCACCACA 400
Qy 301 TTGCTGATGAACTTCCGCGCATGATTTCTATGCTGTTTCATTCGCGCTGATGATGATG 360
Db 401 TTGCTGATGAACTTCCGCGCATGATTTCTATGCTGTTTCATTCGCGCTGATGATGATG 460
Qy 361 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGGCTTACGAGTCACT 420
Db 461 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGGCTTACGAGTCACT 520
Qy 421 GCGGCGACGCGCGCGGCTGATCGGCGGCGGCGGCTTATTCGAGGCTTACGAGTCACT 480
Db 521 GCGGCGACGCGCGCGGCTGATCGGCGGCGGCGGCTTATTCGAGGCTTACGAGTCACT 580
Qy 481 TCCTACTGAGTATTCGCGGCTCTCAACGAGTGGCGATCGGAGTGTATTCCTTTGAA 540
Db 581 TCCTACTGAGTATTCGCGGCTCTCAACGAGTGGCGATCGGAGTGTATTCCTTTGAA 640
Qy 541 ATTAAAGGCTCGAGTTCGCGCTTTCCTCTCTCTTTGTCAGCTGATGATTCCTGCG 600
Db 641 ATTAAAGGCTCGAGTTCGCGCTTTCCTCTCTCTTTGTCAGCTGATGATTCCTGCG 700
Qy 601 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGCTTCAACATTCCTT 660
Db 701 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGCTTCAACATTCCTT 760
Qy 661 GTGTAATTCAGAGTCAGGCGCCCTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGACC 720
Db 761 GTGTAATTCAGAGTCAGGCGCCCTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGACC 820
Qy 721 ATCCGTAATTCCTTCTTGGGAAAGGCTCTAATATGACACTGATTCCTTGGTATTCCTC 780
Db 821 ATCCGTAATTCCTTCTTGGGAAAGGCTCTAATATGACACTGATTCCTTGGTATTCCTC 880
Qy 781 TTGTTGTCAGATATGAGTCAATTAATTTGCGCTTCGCGGCGGTTCCGTTTAAATCC 840
Db 881 TTGTTGTCAGATATGAGTCAATTAATTTGCGCTTCGCGGCGGTTCCGTTTAAATCC 940
Qy 841 TTAAGCCCTCAGTGAATCAATTTGAGGCAAAATGGCATGTGATGTCAGACAGAA 900
Db 941 TTAAGCCCTCAGTGAATCAATTTGAGGCAAAATGGCATGTGATGTCAGACAGAA 1000
Qy 901 TCCTTGCATTTTTCAGCGCATCAACGTTTGCAGCAATGCGATGATGTAAGCTCTAA 960
Db 1001 TCCTTGCATTTTTCAGCGCATCAACGTTTGCAGCAATGCGATGATGTAAGCTCTAA 1060
Qy 961 CCTTGTGCTAATGCGGCTGATTAAGTGTGGGCGATCTTCTTGGGCGGTCAGAGCA 1020
Db 1061 CCTTGTGCTAATGCGGCTGATTAAGTGTGGGCGATCTTCTTGGGCGGTCAGAGCA 1120
Qy 1021 CCTTGTGAGGCTTGGCGCTGAGCAACGTTTGTGTAAGTGTGATCTTTTC 1076
Db 1121 CCTTGTGAGGCTTGGCGCTGAGCAACGTTTGTGTAAGTGTGATCTTTTC 1176

```

```

RESULT 2
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTHAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS

```

```

; CURRENT APPLICATION NUMBER: US/09/471, 803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; OTHER INFORMATION: ATCC13032
; US-09-471-803A-6

```

```

Query Match 100.0%; Score 1076; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGGCTGTGCGCATTCAGAGGACCCCTG 60
Db 101 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGGCTGTGCGCATTCAGAGGACCCCTG 160
Qy 61 GAACGATGATTAAGGTTATCGGCGCTACGAAATCGGCAAGGCTCTAAACCTCCCT 120
Db 161 GAACGATGATTAAGGTTATCGGCGCTACGAAATCGGCAAGGCTCTAAACCTCCCT 220
Qy 121 GCTGAGGTTTGGGCGATGATCCGATTTGGTATTTGGTCTTGGTATTGAAATAC 180
Db 221 GCTGAGGTTTGGGCGATGATCCGATTTGGTATTTGGTCTTGGTATTGAAATAC 280
Qy 181 GGTACGAATGATGAGGAGAGCCCGCATGTTTCGGGCTGATTTTCGGGCGTCCAGCA 240
Db 281 GGTACGAATGATGAGGAGAGCCCGCATGTTTCGGGCTGATTTTCGGGCGTCCAGCA 340
Qy 241 ATGCTGATATGCGCTGCTGTTGAGCGAGCGCCCTGAGGCGCATGCGCTCAGCA 300
Db 341 ATGCTGATATGCGCTGCTGTTGAGCGAGCGCCCTGAGGCGCATGCGCTCAGCA 400
Qy 301 TTGCTGATGATTCGCGCGCATGATTTATGCGTTTCAATTCGCGCTGATGATGATG 360
Db 401 TTGCTGATGATTCGCGCGCATGATTTATGCGTTTCAATTCGCGCTGATGATGATG 460
Qy 361 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGGCTTACGAGTCACT 420
Db 461 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGGCTTACGAGTCACT 520
Qy 421 GCGGCGACGCGCGCGGCTGATCGGCGGCGGCGGCTTATTCGAGGCTTACGAGTCACT 480
Db 521 GCGGCGACGCGCGCGGCTGATCGGCGGCGGCGGCTTATTCGAGGCTTACGAGTCACT 580
Qy 481 TCCTACTGAGTATTCGCGGCTCTCAACGAGTGGCGATGCGAGTGTATTCCTTTGAA 540
Db 581 TCCTACTGAGTATTCGCGGCTCTCAACGAGTGGCGATGCGAGTGTATTCCTTTGAA 640
Qy 541 ATTAAAGGCTCGAGTTCGCGCTTTCCTCTCTTGTGTCAGCTGATGATTCGCG 600
Db 641 ATTAAAGGCTCGAGTTCGCGCTTTCCTCTCTTGTGTCAGCTGATGATTCGCG 700
Qy 601 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGGTTGAGCTTCAACATTCCT 660
Db 701 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGGTTGAGCTTCAACATTCCT 760
Qy 661 GTGTAATTCAGAGTACGCGCTTATTTGGGCGGCTGATTCCTTGGTCTGTGAGC 720
Db 761 GTGTAATTCAGAGTACGCGCTTATTTGGGCGGCTGATTCCTTGGTCTGTGAGC 820
Qy 721 ATCCGTAATTCCTTCTTGGGAAAGGCTGTAATATGACACTGATTCCTCTGATTCCTC 780

```

```

Db      821 ATCCGTAATCTTCTTGGAGAGGCTGCTAAATGACAATCTTCTCTGTAATCTCC 880
QY      781 TTGTTGCGAGATATGCGAGTCAATTAATTTGCGCTCGGAGCGGTTCCGTTTAATCC 840
Db      881 TTGTTGCGAGATATGCGAGTCAATTAATTTGCGCTCGGAGCGGTTCCGTTTAATCC 940
QY      841 TTAAGCCCCACAGATATCAATTTGCGCAAAATGCGAGATGAGTGAAGCGCAGGAA 900
Db      941 TTAAGCCCCACAGATATCAATTTGCGCAAAATGCGAGATGAGTGAAGCGCAGGAA 1000
QY      901 TCCATGCAATTTGACCGCATCAACGTTTGCGAGCAATGCAATGATCTGAACATCTAA 960
Db      1001 TCCATGCAATTTGACCGCATCAACGTTTGCGAGCAATGCAATGATCTGAACATCTAA 1060
QY      961 CCTTTGATCATTTGCGGCTGCGATTAAGTGTGGGCGATCTTTGGGCGTCAGCGA 1020
Db      1061 CCTTTGATCATTTGCGGCTGCGATTAAGTGTGGGCGATCTTTGGGCGTCAGCGA 1120
QY      1021 CCTTTGAGCGTTGGCGTGGACCATCGTTTTGTTGTAAGTGAATCTTTTC 1076
Db      1121 CCTTTGAGCGTTGGCGTGGACCATCGTTTTGTTGTAAGTGAATCTTTTC 1176

```

RESULT 3

US-09-471-803A-2

Sequence 2, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMAN

APPLICANT: EGGELING, LOTMAR

APPLICANT: PEEFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF

FILE REFERENCE: 21123/265496/MAS

CURRENT APPLICATION NUMBER: US/09/471,803A

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 753

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(753)

OTHER INFORMATION: b7mf

OTHER INFORMATION: ATCC14752

US-09-471-803A-2

Query Match 70.0%; Score 753; DB 4; Length 753;
 Best Local Similarity 100.0%; Pred. No. 9.6e-236;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCCATCCAGAGCCCTG 60
Db      1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCCATCCAGAGCCCTG 60
QY      61 GAACAGATGATAAGGTTATCGGCGTACAGAAATCGCGCAAGGCTTAAAAAAGCTCCCT 120
Db      61 GAACAGATGATAAGGTTATCGGCGTACAGAAATCGCGCAAGGCTTAAAAAAGCTCCCT 120
QY      121 GCTGAGGTTTGGCGATGACCCGATTTGTAATGCGTTGCTCTTGTTATTCATATAC 180
Db      121 GCTGAGGTTTGGCGATGACCCGATTTGTAATGCGTTGCTCTTGTTATTCATATAC 180
QY      181 GGTGACGATGATGCGGCGAGCCCACTGTTTCGGGCGTGAATTTTCGGGCGTCCAGCGAA 240
Db      181 GGTGACGATGATGCGGCGAGCCCACTGTTTCGGGCGTGAATTTTCGGGCGTCCAGCGAA 240

```

```

QY      241 ATGCTGATCATGCGCTCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGGCTCAGACA 300
Db      241 ATGCTGATCATGCGCTCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGGCTCAGACA 300
QY      301 TTGCTGTAAGATTCGCGCAGATATCTATGATCGTTTCATTCGCGTGCATGATGATCA 360
Db      301 TTGCTGTAAGATTCGCGCAGATATCTATGATCGTTTCATTCGCGTGCATGATGATCA 360
QY      361 AACCAGATGCGCGTTTCTATTCGTTTTCGCGCTTATTCAGCAAGAGCTTACGATCACT 420
Db      361 AACCAGATGCGCGTTTCTATTCGTTTTCGCGCTTATTCAGCAAGAGCTTACGATCACT 420
QY      421 GCGGCGAGCGCGCAGGCTGATGCGCGTGCAGCTTATCTCATGCAATATAGCGTTTAC 480
Db      421 GCGGCGAGCGCGCAGGCTGATGCGCGTGCAGCTTATCTCATGCAATATAGCGTTTAC 480
QY      481 TCTTACTGAGTATTCGCGGCTCTCACCGAGTGGCGATCGCAGATGATTCCTTTTGA 540
Db      481 TCTTACTGAGTATTCGCGGCTCTCACCGAGTGGCGATCGCAGATGATTCCTTTTGA 540
QY      541 ATTAAGGCGCTGAGTGGCGCTTGTCTCTCTTGTGCAAGTGAATTTGATTCCTGAC 600
Db      541 ATTAAGGCGCTGAGTGGCGCTTGTCTCTCTTGTGCAAGTGAATTTGATTCCTGAC 600
QY      601 CGAAGAAAACAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTACCATTTGCTCT 660
Db      601 CGAAGAAAACAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTACCATTTGCTCT 660
QY      661 GTGTAATTCAGGTCAGGCGCTTATTTGGCGGCTGATCTTGGGCTGTGAGCC 720
Db      661 GTGTAATTCAGGTCAGGCGCTTATTTGGCGGCTGATCTTGGGCTGTGAGCC 720
QY      721 ATCCGTAATCTTCTTTCGGAAGGCTGCTAA 753
Db      721 ATCCGTAATCTTCTTTCGGAAGGCTGCTAA 753

```

RESULT 4

US-09-471-803A-4

Sequence 4, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMAN

APPLICANT: EGGELING, LOTMAR

APPLICANT: PEEFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF

FILE REFERENCE: 21123/265496/MAS

CURRENT APPLICATION NUMBER: US/09/471,803A

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 324

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(324)

OTHER INFORMATION: b7mf

OTHER INFORMATION: ATCC14752

US-09-471-803A-4

Query Match 30.1%; Score 324; DB 4; Length 324;
 Best Local Similarity 100.0%; Pred. No. 8.1e-96;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      753 ATGACAACTGATTTCTCTGTAATCTCTCTTGTGTCAGATATGTCAGTACTTACTTT 812

```

Db	1	ATGACAACTGATTTCTCCGTATTTCTCCTGTGTGTGTGGAGATGTGAGATCTTACTTT	60
QY	813	GCGCCTCGGAGCGGTTCCGTTCTTAAATCCTTAAGCCCCTACGTGAATACAAATTTGGGGC	872
Db	61	GCGCCTCGGAGCGGTTCCGTTCTTAAATCCTTAAGCCCCTACGTGAATACAAATTTGGGGC	120
QY	873	AAAATGGCAGTGTGATGTGCACAGAGAAATCCCTGCACTTTTGAACGAGATCAAGCTTCCG	932
Db	121	AAAATGGCAGTGTGATGTGCACAGAGAAATCCCTGCACTTTTGAACGAGATCAAGCTTCCG	180
QY	933	AGCAATGCGATAGATGTGAAGACTTAACCTTTGGTCTCATTTGCCGTGGAGATTACAGTG	992
Db	181	AGCAATGCGATAGATGTGAAGACTTAACCTTTGGTCTCATTTGCCGTGGAGATTACAGTG	240
QY	993	GTTGGCGCATTTTGTGGCGGTGACGCACTTGTGAACGTTGGGCGCTGGGACCATTCGTT	105
Db	241	GTTGGCGCATTTTGTGGCGGTGACGCACTTGTGAACGTTGGGCGCTGGGACCATTCGTT	300
QY	1053	TTTGTGGACTGGTGAATCTTTTC	1076
Db	301	TTTGTGGACTGGTGAATCTTTTC	324

```

RESULT 5
US-09-489-039A-3016
; Sequence 3016, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Bretton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709, 2004001
; CURRENT FILING DATE: 2000-01-27
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3016
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

```

Query Match	Similarity	Score	58.6	DB	4	Length	756
Best Local	Similarity	52.2%	Pred. No.	6	9e-09		
Matches	130	Conservative	0	Mismatches	119	Indels	0
						Gaps	0
QY	95	TCGCGCAGGCTCTAAAACCTCCCTTGCTGCAAGTTGGGCGATGTACCCGATTCGTATTG	154				
Db	62	TCGGCGCAGGCGCTCAAGAAGACAGTCTCCCATGTCATCAGTACCTTCGGGCGCATTCG	121				
QY	155	CGTTGGTCTCTGGTTATTCAATACGGCTACGAATGATGGGCGAGCCCATGTTTCCG	214				
Db	122	CCTTGGGATTCATAGCTTACCCGGCTGGGCTTACCCCGCGAAGACCTGTCTTCTCT	181				
QY	215	GCCTGATTTTTCGGGGGCTCCACGAAATGCTGTCTATCGCCCTCTGTTGGGGCGACGGC	274				
Db	182	GGATTATTATACCGCCGGCGCAGCAGTATGTATGCACCCGCAATGGTGGCGCGCGCAGTT	241				
QY	275	CCCTGGGCGCCATTCGGCTCACCAACATTCGTGTGAATTCGCCCAAGTATCTATCGT	334				
Db	242	CGCTGTGGGTTCGGCGGCTCGACGGTATGGCGATGTATGCCGACAGTCTATATGGCC	301				
QY	335	TTTCAATCC	343				
Db	302	CTTCACTGC	310				

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367

GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
City: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F18
US-08-232-463-14

Query Match	4.9%	Score 52.4	DB 1	Length 7218
Best Local Similarity	3.6%	Pred.No. 3e-06	Mismatches 157	Indels 0
Matches	14	Conservative 221	Gaps	0
Qy	460	TCATGCAAAATAGCGCTTCACTCCTACTGAGTATTCGCGCGCTCAACCGAGTGGCAGTC	519	
Db	1050	TCGAGGAGAGCTGGATATTT	1109	
Qy	520	GCAAGATTGATCCTTTTGAATTAAAGCGCTCAGATTGCGCCTTTGCTCTCTTTGTC	579	
Db	1110	TT	1169	
Qy	560	ACGCTGACTTGGATTCCTCGCGCAACGAAAGAGATCCCTTCCTGCTGCTCGCAGAT	639	
Db	1170	TT	1229	
Qy	640	TTGAGCTTCACCATGTCTTTGTGTATATCCAGGTCAGGCCATTATTTGCGGCGCTGCTG	699	
Db	1230	TT	1289	
Qy	700	ATCTTCTTGSGTCTGTTGACCATCCGACTCTTTCTTGGAAGAGCTGCTAATGACAA	759	
Db	1290	TT	1349	
Qy	760	CTGATTTCTCTGATATCTCTCTGTGTGCGAGATATGCGACTATCTTTTGCGCTGC	819	
Db	1350	TT	1409	

```
QY      820 GGGCGTTCGTTCTTAATCCTTAAGCCCCCTA 851
      :  ::  ::::  ::::  ||  |
Db     1410 YYYYYYYYYYYYYYYYYYYYYYGTACCAA 1441
```

RESULT 7

US-09-107-532A-1266 / Sequence 1266, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Denise
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (8) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-09-107-532A-1266

QY		287	TCCGCTCACCACATTCGTGTGAACITCCG	317
Db		272	TGGTTTTCTGCACCTTCTAGTCAATTCCG	302

RESULT 8

US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11376
LENGTH: 1038
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11376

RESULT 9

```

US-09-252-991A-11418
/ Sequence 11418, Application US/09252991A
/ Patent No. 6651795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 11418
/ LENGTH: 1086
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11418

```


Db 222 TGTCCGCCATCGCTTCGCGCGCGCGCCCACTGTGTGCGCATCGGCATGCTCAAGGCG 281
Qy 269 CAGCGCCCTGGAGCGCGCATCGCGCTGACACCATTTGCTGGAATTCGCGCAGTATTCT 328
Db 282 GAGCGCGGTTCTTTGATCATCTTACACACCTCTCTGACCTGCGACGACCTGCTCT 341
Qy 329 ATGCGTTTTCATTCGCGCTG 348
Db 342 ACGGATGAGCTTGGCGCG 361

RESULT 10
US-09-621-976-8976/c
Sequence 8976, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.6%; Score 39; DB 4; Length 399;
Best Local Similarity 12.6%; Pred. No. 0.012;
Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

Qy 509 GAGTGGGATCGGACGAGTTATCTTTGAAATTAAAGGCGCTCGAGTCCGCTTGTCT 568
Db 242 SAAMRRKMTKMGSRSGSTGYAAMMYKMSCTSRKMYKRRKRRCTSTKRT 183
Qy 569 CTCTCTTTCAGCTGACTTGTGATCTCTGCGAAGAAAAGACAGATCCCTTCTCTGC 628
Db 182 CYGSGTCKMAYTKRKRMTWTYTYYSVMKMKRMKAYTWTBKMKTTRTKWTM 123
Qy 629 TGCTCCGAGCTTGAGCTTACCATTTGCTTGTGTAATTCAGGCTCAGGCTTTTG 688
Db 122 CTTCWCKCTTVMAGTMMRYRYYAKRAKMSKRCSTWSTCYCMKYNAKKMSYMSMS 63
Qy 689 CGGCGCTGCTGATCTTCTGCTGTGTGACCATCGGCTACTCTTCTTGGAAGAGCT 747
Db 62 MMKMGSMKMYTYTYTYMMKMSKMTYMSMCAKCKRYTAKTYTWTCTWTGKRT 4

RESULT 11
US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: m1ec-feature
LOCATION: 16

OTHER INFORMATION: n=a, y, c or t
US-09-621-976-18033

Query Match 3.4%; Score 37; DB 4; Length 474;
Best Local Similarity 14.4%; Pred. No. 0.059;
Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

Qy 422 CGGCGAGCGCGGAGCGGCTGCTCGCGCTGAGCTTATTCGAAATAGGCTTCACT 481
Db 13 YSGNCTCYCMARKRSTYKGRMYKMSRGSSCYKSCMCMMSRYCGSYKTTTWTW 72
Qy 482 CCTACTGGATTTTCGCGCTGACCGGAGTGCAGTGCAGATGATTCCTTTGAA 541
Db 73 WTTTTCGKABRRMSGGGKTTVMGSKKTKSCMGRWKGYYRMYCYKACYYW 132
Qy 542 TTAAGGCGCTCGAGTTCGCTTGTCTCTCTTTGTCAGGCTGACTTGGATTCGCG 601
Db 133 WKRWYSSCCMWTYKGGSMWTTTMMRRRKYKWTGKKKKTMMMAACYTWR 192
Qy 602 GAAGAAAAGACAGATCCCTTCTCTGCTGCTGCGAGCTTGAAGCTTCAAGCTTCTTG 661
Db 193 YMMMRRAAAKATYTYCMKSKMCCMACCMCMRABASCCMSBMSYTYMCCYY 252
Qy 662 TGATAATTCAGTCAAGCCCTATTTGCGCGCTGCTGATCTTGTGGCTGTGAACA 721
Db 253 MMYGGMYYMMRWGMKEMMYMKKSMWKSMMWRAMWARKTYTYTAMYYTYTYKR 312
Qy 722 TCCGTAATCTTCTTGGGAAAGCTGCTAATGACACT 761
Db 313 MCTYMRKTYTMMMYSRMWSMTARGAAMWCMYMY 352

RESULT 12
US-09-252-991A-3205/c
Sequence 3205, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3205
LENGTH: 1140
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3205

Query Match 3.4%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.12;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 211 TCCGCGCTGATTTTCGCGGCTCCACCGAATGCTGTCATCGGCGCTTGTGGGCGCA 270
Db 465 TACTTCTGATCAACTTCGCGCGGCTCGCGCGCTGATCGGCTGACTTCTGGGCTC 406
Qy 271 GCGCCCTGGCGGCGCATGCGCTCACCATTTGCTGTGAATTCGCGCAAG 322
Db 405 GCGCGCGAGCGGCGACTTCTGTGTACCGCGCTGTCTCTGTAG 354

RESULT 13
US-09-252-991A-2761
Sequence 2761, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

உயிர்ப்பிழை

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAP

```

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/417,884A
FILING DATE: 17-Apr-2003

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051,571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinella, Pamela Deneké
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-10-417-884A-1266

Query Match      3.7%; Score 39.8; DB 6; Length 768;
Best Local Similarity 49.3%; Pred. No. 0.0085;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

```

107 TAAAACTCCCTGCTGCGAGCTTGGGCAATGACCCGATTGATGCGTTGGCTCT 166
 167 TGGTATCAATAGGCTAGCAATGATGGGAGCCCACTTTTCCGCGCTGATTTCC 226
 152 TTGGGAAGCTGCCGAGTTTCAATCCACTAGTCGTCAGTTGATGCCCTGCTGATG 211
 227 CGGGCTCCAGGAATGCTGTCATCGCCCTGTTGGGCGCAGCCGCCCTGGAGCCCA 286
 212 CTGGTTCTGCCCAATTATTCACAGTCAGCAGCTGCTGATGACAGCCCATTTCTTCCA 271
 287 TTGGGCTCAGCAATGCTGTTGAACCTTCCG 317
 272 TCGTTTCTCGACCTTTCTAATCAATTCCTCG 302

Db

RESULT 3
 US-10-767-701-5263
 Sequence 5263, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.

```

APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5263
LENGTH: 619
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS11974_1
US-10-767-701-5263

Query Match      3.3%; Score 35.4; DB 6; Length 619;
Best Local Similarity 54.1%; Pred. No. 0.21;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

226 GCGGCTCCACCGAATGCTGTCATGCGCCCTGTTGTGAGCGAGCGCCCTGGAGGCC 285
 132 GCGGCTTCACCTTGCTGCTGCTGTCGAGCGCTGCGGACAGTGTGCGCTC 191
 286 ATCGGCTCAGCAGATGCTGTAATTCGCGCAGTATTCATTTCTTCCG 345
 192 ATCGGCGAGCGGCTGCTGATCTTACGCCCCGCTGCTGTCGCGGCGCATCAG 251
 346 CTGCATGTGTCA 358
 252 GTGGCGCTGCTCA 264

Db

RESULT 4
 US-10-767-701-10866
 Sequence 10866, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 FILE REFERENCE: 38-21(53535)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 10866
 LENGTH: 1034
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
 US-10-767-701-10866

```

Query Match      3.2%; Score 34.6; DB 6; Length 1034;
Best Local Similarity 49.7%; Pred. No. 0.52;
Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

241 ATGCTGTCATGCGCCCTGTTGTGGCGCAGCCCTCGGCGCATGCGCTACACA 300
 151 ATGGCGGCCCTTACTAGCGCGAGGCCCTGCGCCACAGCGCGCGCGCCCT 210
 301 TTGCTGTGAATTCGCGCAGTATTCATGCTTTTCAATTCGCGTGCATGTGTA 360
 211 GTGGCGGTGTGTCGCCGAGTTCGCGCGCGTTCAGTGTGCGCTTACGTTGAC 270
 361 AACCCATGCGCGTTCATTCGCTTTCGCGTTCAGAGAGCTTACGCACTC 417
 271 AAGGCATAGCTTCTCCAGCGGCACTTCAACCTCACCGAGCCAGCGCGCGTC 327

Db

RESULT 5
 US-10-100-683-983/c

Sequence 983, Application US/10100683
GENERAL INFORMATION:
APPLICANT: Rosen, et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS900
CURRENT APPLICATION NUMBER: US/10/100,683
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: US 60/040,162
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: US 60/043,576
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,601
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,314
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,632
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,892
PRIOR FILING DATE: 1997-08-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 13468
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 983
LENGTH: 3173
TYPE: DNA
ORGANISM: Homo sapiens
US-10-100-683-983

Query Match 3.2%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.99; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGTCATCGCCCTCTGTGTGGCGCAGCCCTCGGCGCCATCGCCCTCACACCA 300
DB 2360 CAGCTCCCGCCCGCCATGTGTGGCTCAGGCGTCTCTGGGGCCATGGGTCTACCCACA 2304

RESULT 6
US-10-100-683-11511
Sequence 11511, Application US/10100683
GENERAL INFORMATION:
APPLICANT: Rosen, et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS900
CURRENT APPLICATION NUMBER: US/10/100,683
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: US 60/040,162
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: US 60/043,576
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,601
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,314
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,632
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,892

PRIOR FILING DATE: 1997-08-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 13468
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 11511
LENGTH: 7071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-100-683-11511

Query Match 3.2%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.6; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGTCATCGCCCTCTGTGTGGCGCAGCCCTCGGCGCCATCGCCCTCACACCA 300
DB 785 CAGCTCCCGCCCGCCATGTGTGGCTCAGGCGTCTCTGGGGCCATGGGTCTACCCACA 841

RESULT 7
US-10-767-701-7545
Sequence 7545, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Zhou, Yihua
APPLICANT: Kowalic, David K.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 7545
LENGTH: 666
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS29292_1
US-10-767-701-7545

Query Match 3.2%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.47; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 215 GCCGATTTTGGCGGGGCTCACCGGAATGCTGTGATCGCCCTGTGTGGCGCAGCG 274
DB 31 GCCGACCAACAGAGGAGAGAGACGCCGTCTCAACCCGTTCTGTGAGCGGCCCGC 90
QY 275 CCTGGGCGCCATCGCGCTACCAACATGTGTGAACTTCCGCCAG 322
DB 91 CGCTGGGCTCCCTGGCGTGCAGCGCGGTCTGTGCTGTGGCGAG 138

RESULT 8
US-10-767-701-744/c
Sequence 744, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kowalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 744
LENGTH: 561
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
US-10-767-701-744

Page 4

Query Match	3.1%	Score 33.8	DB 7	Length 24477
Best Local Similarity		Pred. No. 5		
Matches 68, Conservative				
			Mismatches 0	Gaps 0

302 TGGTGGGAACCTTCCGCCACGATTCTATGCGTTTCATCTCCGCGCGCATGTGTCACAAA 361
37 17863 GCTTCAGAGCACC CGGACACATTTCACTTCATCGATGCGCTTCCCGCTGCTGGGATTTCTACC 17722

Qy	362	ACCCCATG	372
Db	17723	TGCCCCTG	17733

RESULT 11
US-09-804-291A-500
: Sequence 500, Application US/09804291A

```

1 TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
2
3 FILE REFERENCE: 100337.54287US
4
5 CURRENT APPLICATION NUMBER: US/09/804,291A
6
7 PRIOR APPLICATION NUMBER: 2001-03-13
8
9 PRIOR FILING DATE: 2000-03-13
10
11 PRIOR APPLICATION NUMBER: 60/188,914
12
13 PRIOR FILING DATE: 2000-03-13
14
15 PRIOR APPLICATION NUMBER: 60/192,033
16
17 PRIOR FILING DATE: 2000-03-24
18
19 PRIOR APPLICATION NUMBER: 60/198,474
20
21 PRIOR FILING DATE: 2000-04-14
22
23 PRIOR APPLICATION NUMBER: 60/199,335
24
25 PRIOR FILING DATE: 2000-04-24
26
27 PRIOR APPLICATION NUMBER: 60/207,702
28
29 PRIOR FILING DATE: 2000-05-26
30
31 PRIOR APPLICATION NUMBER: 60/213,849
32
33 PRIOR FILING DATE: 2000-06-23
34
35 PRIOR APPLICATION NUMBER: 60/226,534
36
37 PRIOR FILING DATE: 2000-08-16
38
39 PRIOR APPLICATION NUMBER: 60/330,732
40
41 PRIOR FILING DATE: 2000-09-07
42
43 PRIOR APPLICATION NUMBER: 60/266,862
44
45 PRIOR FILING DATE: 2001-02-07
46
47 NUMBER OF SEQ ID NOS: 531
48
49 SOFTWARE: Patentln Ver. 3.2

```

```

; ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match      3.1%
Score 33.2; DB 5; Length 930;

```

QY	467	AAATGCGTTTACCTCCTACCTGGTAATTCGGCGGCTCA	CGGAGATGGGAGATCGAGAGCT	526
Db	509	ACATTCGTGATACCTATTA	CTGTGACTATGGGCCCTCTCGTGAGCTGCTGCTCAGACA	568
QY	527	TCATTCCTTTGAATTAAGGCGCCGAGTTGGCCCTTGCTCTCTCTTCACGCTGA		586
Db	569	CAAGCCTCTTAAGAACTGATGTCATCTCTTGCCCGCTGTACTCTCATGATTA	CTCTGG	628
QY	587	CTTTGG	592	
Db	629	TGCTGG	634	

Tue Apr 20 06:47:15 2004

us-10-608-504-1_copy_101_1176.rnpn

Page 6

Db 746 ATGACGGTGGCCCTGCTTGCGTGTGCTGTCTTGTGTCGACCATTTGACACCCGATGGT 805
QY 669 TCCAGATCAGAGCCCTATTTTGGGCGCTGTATCTTCTTGGAGTCGTGACCATCCGGTA 728
Db 806 GCCCCACATGACGCCCACTGCGCGGTGTCTGTGTGCTGTGGCTGCGTGGGCCATCGTGA 865
QY 729 CTTC 732
Db 866 GTCC 869

Search completed: April 17, 2004, 20:12:57
Job time : 126.34 secs

Tue Apr 20 06:47:15 2004

us-10-608-504-1_copy_101_1176.rst

Page 2

Email: Raf.Podowski@cgr.ki.se
Class: NotI site.

FEATURES
source location/Qualifiers

1..473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human NotI clones"

ORIGIN

Query Match 5.1%; Score 54.6; DB 28; Length 473;
Best Local Similarity 56.2%; Pred. No. 0.0009; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 77;

193 TGGGACAGCCCACTGTTTCCGCGCTGATTTTCGCGGAGTCCAGCAATGCTGTGATC 252
283 TGGAGAGGCGAGGAGGCTCTCGGCGATCGTTCGCGGAGGCGCGAGTGTGCGCATC 342
253 GCGCTGCTGTGGGCGGAGCGCCCTGGGCGCATGCGCTCAGCAATGCTGTGATC 312
343 GGCATGCTTAAAGCGAGCCCACTGCTGATCTGCTACCACTGCTGTGACT 402
313 TTCGCGACGATTTATGCGTTTTCATTCGCGCTGATGCTGCAAAACCCCAT 368
403 TCGACACACTGCTTACGCGCTTGTCCATGCGCGCTTGTGACAGGACCCCTT 458

RESULT 2
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1015B03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID: CSOD1015CA02NP1.

FEATURES
source location/Qualifiers

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="CSOD1015B03"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 4.2%; Score 45; DB 13; Length 1201;
Best Local Similarity 12.9%; Pred. No. 0.65;
Matches 47; Conservative 168; Mismatches 188; Indels 0; Gaps 0;

207 GTTTCGCGGCTGATTTTCGCGGCTCAGCAATGCTGTGATGCGCTGCTGTGAG 266
745 STBTSTSTBTSTBTSTBTSTBTSTBTSTBTSTBTSTBTSTBTSTBTSTBTST 804

267 CGAGCGCCCTGGGCGGCGCTGACCACTGCTGTGATGCTGCGGACGATTT 326
805 SBTCSSSSSSSBTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 864

327 CTATGCGTTTCAATCCCGCTGATGTCGCAAAAACCCCATTTCCGCTTCAATCGGT 386
865 YSSBSST 924
387 TTTCGCGCTTATGACAGCAAGCTTACGAGTCTGCGGAGGCGCGGAGCTGTGAGC 446
925 BTSSSBST 984
447 GTGCGCACTATCTCAATGCAAAATAGCTTCACTCTGCTGCTGCTGCTGCTG 506
985 VTSSBTTST 1044
507 CGAGTGGCGATCGCAGATGATTCCTTTGAAATTAAGGCGCTGAGTTCGCGCTTG 566
1045 SSSSSSSSSSSSSSBTTTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1104

567 CTC 569
1105 TTS 1107

RESULT 3
BX381961 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1072YF05 3-PRIME, mRNA sequence.
ACCESSION BX381961
VERSION BX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID: CSOD1072CC03NP1.

FEATURES
source location/Qualifiers

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="CSOD1072YF05"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.8%; Score 40.8; DB 13; Length 1201;
Best Local Similarity 3.0%; Pred. No. 9.4;
Matches 17; Conservative 191; Mismatches 362; Indels 0; Gaps 0;

297 CACATGCTGTGATGATTCGCGCAGTATGATGCTTTCATTCGCGCTGATGCTGTG 356
406 CCGGNNNNNNKK 465
357 CAAAACCCCATGCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCAGTACGAGT 416

```

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0C002YM12"
/clone_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

INRA, UMR INRA-ENSAR Genétique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2 23 48 54 63
Fax: +33 (0) 2 23 48 54 70
Email: Madeleine.Douaire@roazhon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at siduenasupport@ouv.inra.fr to obtain the chromatogram of this

sequence.
Plate: 0001 row: d column: 10
Seq primer: M13R.
Location/Qualifiers

FEATURES

source

1..692
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gca0001b.d.10"
/tissue_type="adipose tissue, granulosa, multi-tissues,
oviduct, small follicle, utero-vaginal gland"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="AGENAE Gallus gallus multi-tissues library
(gca0)"
/note="Vector: pT773D-pac; Clone distribution: AGENAE
Resource centre, Francois Piumi, Francois Piumi, Inra.fr,
INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine
de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 3.7%; Score 39.4; DB 13; Length 692;
Best Local Similarity 62.9%; Pred. No. 17;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 705 CTTGGGCTGTTGACCATCCGCTACTTCTTCTGGGAAAGGCTGCTAATGCAACTGAT 764
Db 321 CTTGGGCGAGTGGCCATCTGCTCTGCGACCTTGCACAGACGACGTGACCGCAT 262
QY 765 TTCTCCGTATCTCTCTGTTGTGCGAGATGTCAG 801
Db 261 TTTTGCTAAGTCACTTCTGCTGCGAGCTGTCAG 225

RESULT 6 718 bp mRNA linear EST 25-JAN-2001
AJ397466 dke2426 Gallus gallus cDNA clone 217r1, mRNA sequence.
LOCUS AJ397466
DEFINITION AJ397466
ACCESSION AJ397466
VERSION AJ397466.1 GI:7130433
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 718)
Abdrakhmanov, I., Lodygin, D., Gerolt, P., Arakawa, H., Lav, A.,
Plachy, J., Korn, B. and Buerstedde, J.M.
A large database of chicken Bursal ESTs as a resource for the
analysis of vertebrate gene function
Genome Res. 10 (12), 2062-2069 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..718
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="217r1"
/tissue_type="Bursa of Fabricius"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/clone_1lb="dke2426"

FEATURES

source

ORIGIN
Query Match 3.7%; Score 39.4; DB 9; Length 718;

Best Local Similarity 62.9%; Pred. No. 18;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 705 CTTGGGCTGTTGACCATCCGCTACTTCTTCTGGGAAAGGCTGCTAATGCAACTGAT 764
Db 117 CTTGGGCGAGTGGCCATCTGCTCTGCGACCTTGCACAGACGACGTGACCGCAT 58
QY 765 TTCTCCGTATCTCTCTGTTGTGCGAGATGTCAG 801
Db 57 TTTTGCTAAGTCACTTCTGCTGCGAGCTGTCAG 21

RESULT 7 743 bp DNA linear GSS 19-JUN-2003
CC684922
LOCUS CC684922
DEFINITION CC684922 ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0397014,
genomic survey sequence.
ACCESSION CC684922
VERSION CC684922.1 GI:32089698
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 743)
Whitelaw, C.A., Quakenbush, J., Van Aken, S., Uteback, T.,
Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGUX55TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Classes: sheared ends.
Location/Qualifiers
1..743
/organism="Zea mays"
/mol_type="Genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0397014"
/clone_1lb="ZM 0.7.1.5 KB"
/note="Vector: pBCSK-; Site: 1; HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES

source

ORIGIN
Query Match 3.7%; Score 39.4; DB 29; Length 743;
Best Local Similarity 53.6%; Pred. No. 18;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 431 CCGCAGGCTGTCGGCGTGGCGACTTATCTCAATGCAATAGGCTTCACTCTAGTGGG 490
Db 475 CCGCGCGCTCCGCGCGTTCGCTCTCCACATCTCATCTCGCATGGAATATCTAGG 534
QY 491 TATTCGGCGCTCTCACCGAGTGGGATTCGCGAGATTCCTTTGAATTAAGGCGC 550
Db 535 GTTTCGGCGCGCGCACCGCATTCAGCTGGGCTTGGCGAGCGCTCTTAAGCTTAGGTC 594
QY 551 TCGAGTTCGCCCTTGTGCTCTCTCTTGTGTCAGCG 583
Db 595 ACGAGAGGCGCTTGTCTCTCTCTCCGCTGCGCGC 627

RESULT 8 756 bp mRNA linear EST 29-NOV-2002
BU448643/c
LOCUS BU448643
DEFINITION 603211326F1 CSRRAN13 Gallus gallus cDNA clone Chest191c7 5', mRNA

sequence.
 accession B048643 GI:25937954
 version B048643.1
 keywords EST.
 source Gallus gallus (chicken)
 organism Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianidae; Gallus.
 reference 1 (bases 1 to 756)
 authors Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 title A Comprehensive Collection of Chicken cDNAs
 journal Curr. Biol. 12 (22), 1965-1969 (2002)
 medline 2235534
 pubmed 12445392
 comment Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 location/Qualifiers
 source 1..756
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CBEST191c7"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="cSEORN13"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
 EcoRI; Site_2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was bluntly ligated to NotI adaptors, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN
 Query Match 3.7%; Score 39.4; DB 13; Length 756;
 Best Local Similarity 62.9%; Pred. No. 18;
 Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 Y 705 CTGGGTCCTGATGACATCGGACTTCTCTGGGAAGGCTGCTAATGACAATGAT 764
 DB 270 CTTGGGACATGCGCATCTGCTCTGACACTTTGACACAGACGACGACCGCAT 211
 Y 765 TTCTCCTGATTTCTCTGTTGTCGAGATGTCAG 801
 DB 210 TTTTGCTAAGTCACTTCTGCTGACGCTGACAG 174

RESULT 9
 CNS009WA/c 787 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
 DEFINITION BACR20X04 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 accession AL054013
 version AL054013.1 GI:4935590
 keywords GSS.

SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 reference 1 (bases 1 to 787)
 authors Genoscope.
 title Direct Submission
 journal Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 comment Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuhiro Osoegawa and
 Aaron Mammoeer in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; on bw sp. the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 location/Qualifiers
 source 1..787
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR20X04"
 /clone_lib="RPCI-98"
 /note="end : 17"

ORIGIN
 Query Match 3.7%; Score 39.4; DB 29; Length 787;
 Best Local Similarity 45.5%; Pred. No. 18;
 Matches 76; Conservative 15; Mismatches 76; Indels 0; Gaps 0;
 Y 552 CGAGTGGCCCTTGTCTCTCTCTGTCACGCGCTTGTGATTCGCGGACGAAA 611
 DB 334 CTATTGCTCTCTCTCTCTCTCTGTTGAAMKMGKMMCTGGAATCCCTKXGABARAAC 275
 Y 612 GCAATGCCCTTCTCTGTCGTCGACGTTGAGCTTACCATTCCTTGTGTAATCC 671
 DB 274 AYTGTATKTTTCTCYACMGACCTTAAGGTAATTTCTTAATTTTATTTTAA 215
 Y 672 AGGTCAAGCCCTATTTCGCGGCTGATCTTCTGAGTCTGTGTA 718
 DB 214 TAATTGCCATATGATTATATGATTTTATATGTCGCGTGA 168

RESULT 10
 EX425603 885 bp mRNA linear EST 15-MAY-2003
 LOCUS BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 DEFINITION C10B8022A07 3-PRIME, mRNA sequence.
 accession BX425603
 version BX425603.1 GI:30770486
 keywords EST.
 source Homo sapiens (human)
 organism Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 reference 1 (bases 1 to 885)
 authors Li, W.B., Gruber, C., Jessee, J. and Polyses, D.
 title Full-length cDNA libraries and normalization
 journal Unpublished (2001)
 comment Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1b="RPO1-98"
/note="end : TET3"

ORIGIN

Query Match 3.6%; Score 39; DB 29; Length 925;
Best Local Similarity 14.2%; Pred. No. 26;
Matches 37; Conservative 117; Mismatches 106; Indels 0; Gaps 0;

195 GGACAGCCCACTGTTTCCGCGCTGATTTCCGGGCTCCACCAATGCTGTCATGAC 254
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
638 SSSSSSTSSSTSSSTSSSSSSSSSSSSSSSTTSTASGSGWSAGSGSGTSGTSS 697
255 CCTCGTTGGCGAGCGCCCTGCGGCGCATCGCGCTCACCACTGCTGTAATT 314
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
698 SSSSSSTSS 757
315 CCGCCACGATTCATGCGTTTTCATTCGCCGCTGATGTCGCAAAACCCCATGCGCG 374
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
758 GCCTTCCGCTSS 817
375 TTCTATTGCGTTTTCGCGCTTATGACGACGACCTACGACGACTGCGCGCGCGCGC 434
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
818 YAMBCTSTSCGSSSSSSSGKGVTKCGCGCGSSSTNMBGTSSACSSSSSSSSSSSV 877
435 AGGCTGTGCGCGCTGCGCAGC 454
878 SSKSSASSSSSVSSSSSSSVS 897

RESULT 13

BX361080

LOCUS

BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA

DEFINITION

BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA

ACCESSION

BX361080.1 GI:30374504

VERSION

EST.

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 1201)

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope, Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact: Feng Liang Email: filang@lifetech.com URL:<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSOD1079DGB08NP1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="RNA"

/db_xref="taxon:9606"

/clone="CSOD1079YN16"

/issue_type="PLACENTA COT 25-NORMALIZED"

/clone_1b="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was placed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

primer. Five prime end enriched, double-strand cDNA was

primer. Five prime end enriched, double-strand cDNA was

primer. Five prime end enriched, double-strand cDNA was

primer. Five prime end enriched, double-strand cDNA was

ORIGIN

Query Match 3.6%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 10.5%; Pred. No. 38;
Matches 34; Conservative 147; Mismatches 142; Indels 0; Gaps 0;

digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

19 ATTCATTCAGCTGAGAGTGTCCCATCCAGAGCAGCTGAGACCATGATTAAGT 78
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
825 MDDMMMMMMKKNDKKKKNDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 884
79 TATCGCGCTACGAAATCGCGCAAGCTCTAAACCTCGTCTGAGGTTGGCATG 138
885 KKKKKDAKK 944
139 TACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 198
945 KKK 1004
199 GCCCACTGTTTCCGCGCTGATTTTCCGCGCTCCACCAATGCTGTCATGCGC 258
1005 KNDMMMMMMKK 1064
259 GTTGTGCGCGCAGCGCCCTGCGCGCATCGCTACACCATGCTGTAATTCCG 318
1065 KKKKKKKCCCBKCCCKKCKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1124
319 CAGCATTCATGCGTTTCATT 341
1125 CCKKKKKCKKK 1147

RESULT 14

CB657749

LOCUS

CB657749 835 bp mRNA linear EST 09-APR-2003

DEFINITION

OSJNEC13F02.f OSJNEC Oryza sativa (japonica cultivar-group) CDNA

ACCESSION

CB657749.1 GI:29661474

VERSION

EST.

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 835)

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.

Large-scale identification of ESTs involved in the interaction

between rice and Magnaporthe grisea

Unpublished (2003)

Contact: Rod Wang

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: <http://genome.arizona.edu>

PCR Primers

FORWARD: gga aac cga cgg cca gtc

BACKWARD: gga aac agc tac gac cat g

Plate: 13 Row: F Column: 02

Seq primer: gta aac cga cgg cca gtcg.

Location/Qualifiers

1..835

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="RNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEC13F02"

/issue_type="Leaf"

ORIGIN

```

/dev_stage="3 week"
/lab_host="DH10B"
/clone_1ib="OSJunc"
/Note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

```

Query Match 3.5%; Score 38; DB 14; Length 835;

Best Local Similarity 51.1%; Pred. No. 46; Mismatches 85; Indels 0; Gaps 0;

Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

QY 252 CCGCCCTGCTTGGGCGCAGCCCGCTGAGCGCATCGGCTCACCACATTGCTGTGA 311
DB 79 CACACCGGCTCTCGGCGCGCGCTGCTGCTTCCGCGCGCTGCGAGATGCGCTC 138
QY 312 CTTCCGCGCAGCTATTCTATCGCTTTTCATCCCGCTGATGTGTCAAAAACCCATTGC 371
DB 139 ATGCGCGCGCGGTGTCGACCGATCGCGACCCGACCTGCGCGCTGGAAGCGCTGCT 198
QY 372 CCGTTCTATGCGCTTTCGCGCTTATCGACGAGCTACGAGTCACTGCGGC 425
DB 199 CTCGCCCTACGCGCTCTTCGCGCAGGACGCTGCGCTGCGCTGCGCT 252

```

RESULT 15

CNS005NG/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Oosawa and
 Aaron Mammose in Peter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPL-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 pl and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

```

1. .995
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12G04"
/clone_1ib="RPL-98"
/Note="end : TET3"

```

ORIGIN

Query Match 3.5%; Score 37.8; DB 29; Length 995;
 Best Local Similarity 34.0%; Pred. No. 57;
 Matches 54; Conservative 41; Mismatches 64; Indels 0; Gaps 0;

```

QY 916 CCGCATCAAGCTTTCGACGAATGCGATGATCTGAAGACTTAACCTTGTCTCATTTG 975
DB 872 CAGTMTTACGAGCGCTGTGTGTAAGASASATTTGVRSTGCGCTTTMTTATTAAT 813
QY 976 CCGTTCGATTACAGTGTGGCGCATCTTCTGGCGCTGACGACCTTTGAGCGTTG 1035
DB 812 CSKTTSSBTWGRKACSCMTTTSGSWASWTGCTGKKKGAGGAGGAGGCGGAKKSYTB 753
QY 1036 CGCGTGCACACCATCGTTTGTGTTGACTGTGTAATCTT 1074
DB 752 SSGBTGTGKKTSTARKKTTTGCTTGAGSTGDDTAATCTTT 714

```

Search completed: April 17, 2004, 20:01:41
 Job time : 3485.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 3191.51 Seconds
(without alignments)

10226.291 Million cell updates/sec

Title: US-10-608-504-2

Perfect score: 753

Sequence: 1 gtcaaaacgaagat.....tcttggaagagctgctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_hum.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_pac.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hrgo_hum.*
40: em_hrgo_mus.*
41: em_hrgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	AR391954	AR391954 Sequence
2	753	100.0	753	AX120373	AX120373 Sequence
3	753	100.0	753	AX137710	AX137710 Sequence
4	753	100.0	753	BD014992	BD014992 Nucleotide
5	753	100.0	753	BD162490	BD162490 Novel pol
6	753	100.0	1271	AR391953	AR391953 Sequence
7	753	100.0	1271	AR391956	AR391956 Sequence
8	753	100.0	1271	AX137709	AX137709 Sequence
9	753	100.0	1271	AX137714	AX137714 Sequence
10	753	100.0	1271	BD014991	BD014991 Nucleotide
11	753	100.0	1271	BD014994	BD014994 Nucleotide
12	753	100.0	2105	AR454053	AR454053 Coryneb
13	753	100.0	340000	AP005274	AP005274 Coryneb
14	753	100.0	349980	AX120085	AX120085 Sequence
15	273.6	36.3	397	AX137083	AX137083 Sequence
16	273.6	36.3	778	AX137085	AX137085 Sequence
17	149.6	19.9	87340	AP005224	AP005224 Coryneb
18	132.6	17.6	53793	AX123492	AX123492 Sequence
19	127.4	16.9	711	BD165609	BD165609 Novel pol
20	127.4	16.9	309400	AX127153	AX127153 Sequence
21	127.4	16.9	325651	AP005283	AP005283 Coryneb
22	127.4	16.9	715	BD013995	BD013995 Novel nuc
23	85	11.3	715	BD013995	BD013995 Novel nuc
24	59.8	7.9	349926	AX386287	AX386287 Sequence
25	58.6	7.8	756	AE011050	AE011050 Methanosa
26	57.2	7.6	9888	HS2425448	HS2425448 Homo sapi
27	54.4	7.2	856	AE013585	AE013585 Methanosa
28	51.8	6.9	11025	AE016783	AE016783 Pseudomon
29	51.8	6.8	308015	AE017209	AE017209 Geobacter
30	48	6.4	301838	AX655393	AX655393 Sequence
31	47	6.2	2000	EC410307	EC410307 Erythrina C
32	46.6	6.2	10733	AE013695	AE013695 Yersinia
33	46.6	6.2	10733	AE013695	AE013695 Yersinia
34	46.6	6.2	220050	AJ414156	AJ414156 Yersinia
35	46.2	6.1	7218	166494	166494 Sequence 14
36	46.2	6.1	215050	AL646084	AL646084 Ralstonia
37	45	6.1	20342	AE008888	AE008888 Salmonella
38	45.8	6.0	348624	AX640441	AX640441 Bordetella
39	45.4	6.0	348624	AX640441	AX640441 Bordetella
40	44.6	5.9	329709	AP002957	AP002957 Mesorhizo
41	44.2	5.9	208524	AE016925	AE016925 Chromobac
42	44.2	5.9	264050	AL627279	AL627279 Salmonella
43	44.2	5.9	300592	AE016846	AE016846 Salmonella
44	43.4	5.8	346294	AP002999	AP002999 Mesorhizo

ALIGNMENTS

RESULT 1
LOCUS AR391954
DEFINITION Sequence 2 from patent US 6613545.
ACCESSION AR391954
VERSION AR391954.1 GI:40115725
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 753)
AUTHORS Kemmerknecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Parent: US 6613545-A 2 02-SEP-2003;

FEATURES

source

Location/Qualifiers
1..753
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGCCTCAAGGACCCCTG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGCCTCAAGGACCCCTG 60
QY 61 GAACCAAGATGATTAAGTTATCGGCGCTAGAAATGCGGCAAGGTCTAAAACTCCCT 120
DB 61 GAACCAAGATGATTAAGTTATCGGCGCTAGAAATGCGGCAAGGTCTAAAACTCCCT 120
QY 121 GCTCAGAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGGTTATCAATAC 180
DB 121 GCTCAGAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGGTTATCAATAC 180
QY 181 GGTACAGATGATGGGCGACGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
DB 181 GGTACAGATGATGGGCGACGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
QY 241 ATGCTGATCATCGCCCTCGTTGTGGGCGACGCCCTCGGCGCATCGCGCTCACACA 300
DB 241 ATGCTGATCATCGCCCTCGTTGTGGGCGACGCCCTCGGCGCATCGCGCTCACACA 300
QY 301 TTGCTGGTGAATCTCCGCACTGATTCATGCTTTTCAATCCCGCTGATGTTGTTCAA 360
DB 301 TTGCTGGTGAATCTCCGCACTGATTCATGCTTTTCAATCCCGCTGATGTTGTTCAA 360
QY 361 AACCCATTGCGCGCTTCTATTCGATTCGCTTATCGAGCTTATCGAGAGCTACGAGTCACT 420
DB 361 AACCCATTGCGCGCTTCTATTCGATTCGCTTATCGAGCTTATCGAGAGCTACGAGTCACT 420
QY 421 GCGGCCAGGCCCGACGCTGATGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCAC 480
DB 421 GCGGCCAGGCCCGACGCTGATGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCAC 480
QY 481 TCCCTACGAGTATTCGCGCGCTCAACGAGTGGCGATGCGAGTGAATCTCTTTGAA 540
DB 481 TCCCTACGAGTATTCGCGCGCTCAACGAGTGGCGATGCGAGTGAATCTCTTTGAA 540
QY 541 ATTAAGGCGCTCGAGTTCGCTTCTCTTGTCAAGCTGACTTTGATTCCTGC 600
DB 541 ATTAAGGCGCTCGAGTTCGCTTCTCTTGTCAAGCTGACTTTGATTCCTGC 600
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTCTCGAGAGTTGAGCTTCAACATTGCTCT 660
DB 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTCTCGAGAGTTGAGCTTCAACATTGCTCT 660
QY 661 GTGGTAAATTCAGGCTCAGGCTTATTTGCGCGCTGCTGATCTCTTGGGCTGTGACC 720
DB 661 GTGGTAAATTCAGGCTCAGGCTTATTTGCGCGCTGCTGATCTCTTGGGCTGTGACC 720
QY 721 ATCCGTAATCTCTTCTGGAAGAGGCTGCTAAA 753
DB 721 ATCCGTAATCTCTTCTGGAAGAGGCTGCTAAA 753

```

RESULT 2
AX120373 753 bp DNA linear PAT 11-MAY-2001
LOCUS AX120373
DEFINITION Sequence 289 from Patent EP108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE Corynebacterium glutamicum
Corynebacterium glutamicum
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE

Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Tokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
Novel polynucleotides
Patent: EP 1108790-A 289 20-JUN-2001;
JOURNAL KYOWA HAKKO KOGYO CO., LTD. (JP)

FEATURES

source

Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGCCTCAAGGACCCCTG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGCCTCAAGGACCCCTG 60
QY 61 GAACCAAGATGATTAAGTTATCGGCGCTAGAAATGCGGCAAGGTCTAAAACTCCCT 120
DB 61 GAACCAAGATGATTAAGTTATCGGCGCTAGAAATGCGGCAAGGTCTAAAACTCCCT 120
QY 121 GCTCAGAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGGTTATCAATAC 180
DB 121 GCTCAGAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGGTTATCAATAC 180
QY 181 GGTACAGATGATGGGCGACGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
DB 181 GGTACAGATGATGGGCGACGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
QY 241 ATGCTGATCATCGCCCTCGTTGTGGGCGACGCCCTCGGCGCATCGCGCTCACACA 300
DB 241 ATGCTGATCATCGCCCTCGTTGTGGGCGACGCCCTCGGCGCATCGCGCTCACACA 300
QY 301 TTGCTGGTGAATCTCCGCACTGATTCATGCTTTTCAATCCCGCTGATGTTGTTCAA 360
DB 301 TTGCTGGTGAATCTCCGCACTGATTCATGCTTTTCAATCCCGCTGATGTTGTTCAA 360
QY 361 AACCCATTGCGCGCTTCTATTCGATTCGCTTATCGAGCTTATCGAGAGCTACGAGTCACT 420
DB 361 AACCCATTGCGCGCTTCTATTCGATTCGCTTATCGAGCTTATCGAGAGCTACGAGTCACT 420
QY 421 GCGGCCAGGCCCGACGCTGATGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCAC 480
DB 421 GCGGCCAGGCCCGACGCTGATGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCAC 480
QY 481 TCCCTACGAGTATTCGCGCGCTCAACGAGTGGCGATGCGAGTGAATCTCTTTGAA 540
DB 481 TCCCTACGAGTATTCGCGCGCTCAACGAGTGGCGATGCGAGTGAATCTCTTTGAA 540
QY 541 ATTAAGGCGCTCGAGTTCGCTTCTCTTGTCAAGCTGACTTTGATTCCTGC 600
DB 541 ATTAAGGCGCTCGAGTTCGCTTCTCTTGTCAAGCTGACTTTGATTCCTGC 600
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTCTCGAGAGTTGAGCTTCAACATTGCTCT 660
DB 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTCTCGAGAGTTGAGCTTCAACATTGCTCT 660
QY 661 GTGGTAAATTCAGGCTCAGGCTTATTTGCGCGCTGCTGATCTCTTGGGCTGTGACC 720
DB 661 GTGGTAAATTCAGGCTCAGGCTTATTTGCGCGCTGCTGATCTCTTGGGCTGTGACC 720
QY 721 ATCCGTAATCTCTTCTGGAAGAGGCTGCTAAA 753
DB 721 ATCCGTAATCTCTTCTGGAAGAGGCTGCTAAA 753

```

RESULT 3
AX137710

LOCUS AX137710 753 bp DNA linear PAT 30-MAY-2001
 DEFINITION Sequence 2 from Patent EP1096010.
 ACCESSION AX137710
 VERSION AX137710.1 GI:14273889
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 1 Kennerknecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 2 02-MAY-2001;
 Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 Location/Qualifiers
 1..753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1..>753
 /note="unassigned protein product; brnF"
 /codon_start=1
 /translation="1"
 /protein_id="CAC39925.1"
 /db_xref="GI:14273890"
 /db_xref="EMBL:CAC39925"
 /translation="MOKQIEHSLSEVSPSALEPDKRYRYEIAQGLTSLAQL
 GMYPIGAFGLVIOGYEWMWAPLFGSLIFAGSTEMVILVWGAAPLGAITLIL
 VNEFVFAFSPLEHVNPIARFVSPALIDEMAVTAAPAPMSAMRLISQIAH
 SYWVGGLTGAVALHILPREIKGLEFALCSFVTLTDSCTKQIPSLIAGSFTI
 ALVIRPGALFAALLIFGLTIRIFLGRKAK"

CDS
 1 GTCGAAAAAAGCAAGATGATTCATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 60
 1 GTGCAAAAAAGCAAGATGATTCATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 60
 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
 121 GCTGCAAGTTGGGCAATGATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 180
 121 GCTGCAAGTTGGGCAATGATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 180
 121 GCTGCAAGTTGGGCAATGATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 180
 181 GGTACGAATGATGAGGCAAGCCCACTGTTTCCGCGCTAATTTTCGCGGCTCCACCGAA 240
 181 GGTACGAATGATGAGGCAAGCCCACTGTTTCCGCGCTAATTTTCGCGGCTCCACCGAA 240
 241 ATGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 300
 241 ATGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 300
 301 TTGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 360
 301 TTGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 360
 361 AACCCATTGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 420
 361 AACCCATTGCGCTGTTGTGTGGGCGAGCGCCCTGAGGCGCATCGCGCTCCACCA 420
 421 GGGGCGAGCGCGAGCGAGTGTGCGGCGAGCGCGCATTCATCAATGAATAGGTTTAC 480
 421 GGGGCGAGCGCGAGCGAGTGTGCGGCGAGCGCGCATTCATCAATGAATAGGTTTAC 480
 481 TCTTACTGGGATTCGCGGCTCTCACCGAGTGGCGCATCGAGAGTTGATTCCTTTGAA 540

Db 481 TCTTACTGGGATTCGCGGCTCTCACCGAGTGGCGCATCGAGAGTTGATTCCTTTGAA 540
 Qy 541 ATTAAGGCGCTGAGTGTGCGCTGCTCTCTTTGTGACGCTGATTTGATTCCTG 600
 Db 541 ATTAAGGCGCTGAGTGTGCGCTGCTCTCTTTGTGACGCTGATTTGATTCCTG 600
 Qy 601 CGAAGCAAAAGCAAGATTCCTCTGCTGCTGCGAGGTTGAGCTTCAACATGCTCT 660
 Db 601 CGAAGCAAAAGCAAGATTCCTCTGCTGCTGCGAGGTTGAGCTTCAACATGCTCT 660
 Qy 661 GTGCTATTCAGATGAGCGCCCTATTTGGCGGCTGATCTTCTTGGTCTGTAAC 720
 Db 661 GTGCTATTCAGATGAGCGCCCTATTTGGCGGCTGATCTTCTTGGTCTGTAAC 720
 Qy 721 ATCCGATCTTCTTCTTGGGAAAGCTGTAA 753
 Db 721 ATCCGATCTTCTTCTTGGGAAAGCTGTAA 753

RESULT 4
 BD014992 753 bp DNA linear PAT 27-AUG-2002
 LOCUS BD014992
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014992
 VERSION BD014992.1 GI:2255799
 KEYWORDS JP 2001169788-A/2.
 SOURCE unclassified
 ORGANISM unclassified
 unclassified
 unclassified
 1 (bases 1 to 753)
 Kennerknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 2001169788-A 2 26-JUN-2001;
 Degussa HUBS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC14752
 PN JP 2001169788-A/2
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KENNERKNECHT, HERMANN SAHM, LOTHAR EGERING, WALTER PI

JOURNAL
 COMMENT
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
 PC (C12N15/09, C12R1.15), (C12N1/21, C12R1.15), (C12P13/06, C12R1.15),
 PC (C12N15/00, C12R1.15)
 CC brnF
 CC key
 FH key
 FT CDS Location/Qualifiers
 1..753
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 1..753
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 9.8e-185;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAAAGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 60
 Db 1 GTGCAAAAAAGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 60
 Qy 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
 Db 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
 Qy 121 GCTGCAAGTTGGGCAATGATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 180
 Db 121 GCTGCAAGTTGGGCAATGATTCAGCCCTGAGGTGTGCGCATCCAGAGCCCTG 180

```

QY 181 GGCTACGATGATGGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGCTCCACCGAA 240
DB 181 GGCTACGATGATGGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGCTCCACCGAA 240
QY 241 ATGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGGCGCATCGGCGCTACACA 300
DB 241 ATGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGGCGCATCGGCGCTACACA 300
QY 301 TTGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGGCGCATCGGCGCTACACA 360
DB 301 TTGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGGCGCATCGGCGCTACACA 360
QY 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATCGAGAGAGCTTACGAGCTACT 420
DB 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATCGAGAGAGCTTACGAGCTACT 420
QY 421 GGGGCGAGGCGCGAGGCTGTCGCGGTCGAGCTTATCTCAATGCAATGAGCTTTCAC 480
DB 421 GGGGCGAGGCGCGAGGCTGTCGCGGTCGAGCTTATCTCAATGCAATGAGCTTTCAC 480
QY 481 TCTTACTGGGATATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGATGATTCCTTTGAA 540
DB 481 TCTTACTGGGATATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGATGATTCCTTTGAA 540
QY 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGCTTTCAGCTTTCAGCTTTC 600
DB 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGCTTTCAGCTTTCAGCTTTC 600
QY 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCT 660
DB 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCT 660
QY 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
DB 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
QY 721 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTTAA 753
DB 721 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTTAA 753

RESULT 5
LOCUS BD162490 753 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD162490
VERSION BD162490.1 GI:27686248
KEYWORDS JP 2002191370-A/289.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 753)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
TITLES Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
JOURNAL Novel polynucleotide
Patent: JP 2002191370-A 289 09-JUL-2002;
KYOMA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/289
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIRO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
PC 04, C12P13/08,
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/56,
PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, (C12N1/21, C12R1/15),
PC (C12N1/21, C12R1/13), (C12N1/21, C12R1/01), (C12P13/08, C12R1/15),

```

```

FEATURES
Source
Location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:3264"

ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred.No.9.8e-185; Mismatches 0; Gaps 0;
Matches 753; Conservative 0;

QY 1 GTCAAAAAAGCAGAGATTCAATTCAGCTGAGAGTGTCCCATCAAGGAGCCCTG 60
DB 1 GTCAAAAAAGCAGAGATTCAATTCAGCTGAGAGTGTCCCATCAAGGAGCCCTG 60
QY 61 GAACGAGATATTAAGTTATCGGCGCTACGAAATCGCGAAGTCTAATAAATCCCTCT 120
DB 61 GAACGAGATATTAAGTTATCGGCGCTACGAAATCGCGAAGTCTAATAAATCCCTCT 120
QY 121 GCTGCAAGTTGGGCGATGTCACCGGATGATGATGCTGTTGATCTCTGTTATTCATAC 180
DB 121 GCTGCAAGTTGGGCGATGTCACCGGATGATGATGCTGTTGATCTCTGTTATTCATAC 180
QY 181 GGCTACGATGATGGGAGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 240
DB 181 GGCTACGATGATGGGAGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 240
QY 241 ATGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGCGCATTCGCGCTCACACA 300
DB 241 ATGCTGATGATGCGCCCTGTTTGGGGCGAGCGCCCTGGGCGCATTCGCGCTCACACA 300
QY 301 TTGCTGATGATCTTCGCGCGATTCATTCAGCTTTCATTCGCGCTGATGATGATGAT 360
DB 301 TTGCTGATGATCTTCGCGCGATTCATTCAGCTTTCATTCGCGCTGATGATGATGAT 360
QY 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATCGAGAGAGCTTACGAGCTACT 420
DB 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATCGAGAGAGCTTACGAGCTACT 420
QY 421 GGGGCGAGGCGCGAGGCTGTCGCGGTCGAGCTTATCTCAATGCAATGAGCTTTCAC 480
DB 421 GGGGCGAGGCGCGAGGCTGTCGCGGTCGAGCTTATCTCAATGCAATGAGCTTTCAC 480
QY 481 TCTTACTGGGATATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGATGATTCCTTTGAA 540
DB 481 TCTTACTGGGATATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGATGATTCCTTTGAA 540
QY 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTCGCTGCTGATCTTTCAGCTTTCAGCTTTC 600
DB 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTCGCTGCTGATCTTTCAGCTTTCAGCTTTC 600
QY 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCTGCGAGGTTTGAAGCTTTCAGCTTTC 660
DB 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCTGCGAGGTTTGAAGCTTTCAGCTTTC 660
QY 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
DB 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
QY 721 ATCCGTAATCTTCTTGGGAAAGCTGCTTAA 753
DB 721 ATCCGTAATCTTCTTGGGAAAGCTGCTTAA 753

RESULT 6
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003

```

DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1271)
AUTHORS Kernerkecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;
FEATURES Location/Qualifiers
source 1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAGAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGGACGCTG 60
DB 101 GTGCAGAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGGACGCTG 160
QY 61 GAACAGATGATTAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 120
DB 161 GAACAGATGATTAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 220
QY 121 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 180
DB 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
QY 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
DB 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
QY 181 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACGAA 240
DB 281 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACGAA 340
QY 241 ATGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
DB 341 ATGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 400
QY 301 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 360
DB 401 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 460
QY 361 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 420
DB 461 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 520
QY 421 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATACGTTTAC 480
DB 521 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATACGTTTAC 580
QY 481 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 540
DB 581 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 640
QY 541 ATTAAGGCGCTGAGATTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTTCCTTGC 600
DB 641 ATTAAGGCGCTGAGATTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTTCCTTGC 700
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGAGTTTGAAGTTCACATTCCTCT 660
DB 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGAGTTTGAAGTTCACATTCCTCT 760
QY 661 GTGGAATTCAGAGTCAAGGCGCTTATTGCGGCGCTGCTGATCTTCTTGAGCTGTTGAC 720
DB 761 GTGGAATTCAGAGTCAAGGCGCTTATTGCGGCGCTGCTGATCTTCTTGAGCTGTTGAC 820
QY 721 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 753
DB 821 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 853

RESULT 7
LOCUS AR391956 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1271)
AUTHORS Kernerkecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
FEATURES Location/Qualifiers
source 1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAGAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGGACGCTG 60
DB 101 GTGCAGAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGGACGCTG 160
QY 61 GAACAGATGATTAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 120
DB 161 GAACAGATGATTAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 220
QY 121 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 180
DB 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
QY 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
DB 221 GCTGAGGTTTGGGCAATGATCCGATGGTATGGCTTGTCTCTTGATTCAATAC 280
QY 181 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACGAA 240
DB 281 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACGAA 340
QY 241 ATGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
DB 341 ATGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 400
QY 301 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 360
DB 401 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 460
QY 401 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 460
DB 401 TTGCTGATGATGCGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 460
QY 361 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 420
DB 461 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 520
QY 461 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 520
DB 461 AACCCGATGCGCGTTTCTATTGCGTTTTCGCGCTTATGAGAAAGCTACGACGCACT 520
QY 421 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATAGGTTTAC 480
DB 521 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATAGGTTTAC 580
QY 521 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATAGGTTTAC 580
DB 521 GCGGCGAGGCGCGCAGGCTGTGCGGCGTGGCACTTATCTCAATGAAATAGGTTTAC 580
QY 481 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 540
DB 581 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 640
QY 581 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 640
DB 581 TCCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGA 640
QY 541 ATTAAGGCGCTGAGATTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTTCCTTGC 600
DB 641 ATTAAGGCGCTGAGATTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTTCCTTGC 700
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGAGTTTGAAGTTCACATTCCTCT 660
DB 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGAGTTTGAAGTTCACATTCCTCT 760
QY 661 GTGGAATTCAGAGTCAAGGCGCTTATTGCGGCGCTGCTGATCTTCTTGAGCTGTTGAC 720
DB 761 GTGGAATTCAGAGTCAAGGCGCTTATTGCGGCGCTGCTGATCTTCTTGAGCTGTTGAC 820

Db 761 GTGTAATTCAGGCTCAGGCTATTTGGGCGCTGATCTTCTTGAGTGTGACC 820

Qy 721 ATCCGTAATCTTCTTGGGAAAGCTGCTAAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGCTGCTAAA 853

RESULT 8

AX137709 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 1 from Patent EP1096010.

DEFINITION AX137709

ACCESSION AX137709

VERSION AX137709.1 GI:14273886

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 Kernerkecht N., Eggeling L., Sahm H. and Pfeifferle W.
Nucleotide sequences coding for branched-chain amino acids export
protein, method for isolating them and their use
Patent: EP 1096010-A 1 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)

FEATURES

source

1. 1271
Location/Qualifiers

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1178"

/note="ATCC14752"

101. .856

/note="unnamed protein product; brnF"

/codon_start=1

/transl_table=1

/protein_id="CAC39923.1"

/db_xref="GI:14273887"

/db_xref="EMBL:CAC39923"

/translation="MQQTQETHSSLEVSPPSKALBPDDKGYRREYIAQGLKTSIAAGL
GMYPGIAFGLLVQYGYEMWAPLPSGLIFAGSTEMLVIALVVGAAPLGATLTL
VNFHVFYAFSPFLHVVNPIARFYSFALIDBAVYATAPAGMSAMRLISMQIAFH
SVWFGGLGVAIAELIPEIKGLEPALSFTVLTLDSCRTKQIPSLIAGLSFTI
ALVIVPGALFALLIFGLITIRYPLGKAK"

853. 1179

/note="unnamed protein product; brnE"

/codon_start=1

/transl_table=1

/protein_id="CAC39924.1"

/db_xref="GI:14273888"

/db_xref="EMBL:CAC39924"

/translation="MTTDSCTLIYVAVCAVTPARVAPLILKRIEPSOVGTMAW
WMAGLILATITASTFRSNAIDKTLTFGLIAVAITVAALLGGRITLISVAGTIVFV
GLVNL"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-164;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGCTGCCATCCAAAGCAGCCCTG 60

Db 101 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGCTGCCATCCAAAGCAGCCCTG 160

Qy 61 GAACCAATATTAAGGTATTCGCGCTACGAATCGCGAAGGTCTAAAAAACCCTCCCT 120

Db 161 GAACCAATATTAAGGTATTCGCGCTACGAATCGCGAAGGTCTAAAAAACCCTCCCT 220

Qy 121 GCTGCAAGTTGGCATGTACCCGATTTGATTCGCTTTGCTCTTTGTTATTCATAC 180

Db 221 GCTGCAAGTTGGCATGTACCCGATTTGATTCGCTTTGCTCTTTGTTATTCATAC 280

Qy 181 GGCTACGAATGGTGGGAGGCGCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 240

Db 281 GGCTACGAATGGTGGGAGGCGCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 340

Qy 241 ATGCTGATCATGCGCCCTGTTGTTGGGCGAGCGCCCTGAGGCGCATGAGCGCTCAACACA 300

Db 341 ATGCTGATCATGCGCCCTGTTGTTGGGCGAGCGCCCTGAGGCGCATGAGCGCTCAACACA 400

Qy 301 TTGCTGATGACTTCGCGCAGATATTCATGCGTTTTCATTCGCCGTGATGTGCTAAA 360

Db 401 TTGCTGATGACTTCGCGCAGATATTCATGCGTTTTCATTCGCCGTGATGTGCTAAA 460

Qy 361 AACCCATTCGCCGTTTCTATTTGCGTTTTCGCGCTATGAGCAAGCCATGAGCATCT 420

Db 461 AACCCATTCGCCGTTTCTATTTGCGTTTTCGCGCTATGAGCAAGCCATGAGCATCT 520

Qy 421 GCGGCGAGCCCGCAGCGTGTGCGCGTGCAGATTATCTCATGCAATAGCGTTTCA 480

Db 521 GCGGCGAGCCCGCAGCGTGTGCGCGTGCAGATTATCTCATGCAATAGCGTTTCA 580

Qy 481 TCTACTGCGTATTCGCGGCTTCAACCGAGTGGCGATGCGAAGTTGATTCCTTTGAA 540

Db 581 TCTACTGCGTATTCGCGGCTTCAACCGAGTGGCGATGCGAAGTTGATTCCTTTGAA 640

Qy 541 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTTTGTGCAAGCTGATTCGATTCCTGC 600

Db 641 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTTTGTGCAAGCTGATTCGATTCCTGC 700

Qy 601 CGAACAAAAAGCAGATCCCTTCTGCTGCTGCGAGTTTGAAGCTTACCATTTCTT 660

Db 701 CGAACAAAAAGCAGATCCCTTCTGCTGCTGCGAGTTTGAAGCTTACCATTTCTT 760

Qy 661 GTGTAATTCAGGTCAGGCGCTATTTGGGCGGCTGTGATTCCTTGGTGTGAGC 720

Db 761 GTGTAATTCAGGTCAGGCGCTATTTGGGCGGCTGTGATTCCTTGGTGTGAGC 820

Qy 721 ATCCGTAATCTTCTTGGGAAAGCTGCTAAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGCTGCTAAA 853

RESULT 9

AX137714 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 6 from Patent EP1096010.

DEFINITION AX137714

ACCESSION AX137714

VERSION AX137714.1 GI:14273893

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 Kernerkecht N., Eggeling L., Sahm H. and Pfeifferle W.
Nucleotide sequences coding for branched-chain amino acids export
protein, method for isolating them and their use
Patent: EP 1096010-A 6 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)

FEATURES

source

1. 1271
Location/Qualifiers

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1178"

/note="ATCC13032"

101. .856

/note="unnamed protein product; brnF"

/codon_start=1

/transl_table=1

/protein_id="CAC39927.1"

/db_xref="GI:14273894"

/db_xref="EMBL:CAC39927"

/translation="MQQTQETHSSLEVSPPSKALBPDDKGYRREYIAQGLKTSIAAGL
GMYPGIAFGLLVQYGYEMWAPLPSGLIFAGSTEMLVIALVVGAAPLGATLTL
VNFHVFYAFSPFLHVVNPIARFYSFALIDBAVYATAPAGMSAMRLISMQIAFH
SVWFGGLGVAIAELIPEIKGLEPALSFTVLTLDSCRTKQIPSLIAGLSFTI
ALVIVPGALFALLIFGLITIRYPLGKAK"

CDS

CDS
853. .1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=11
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="REMBL:CA039928"
/translation="MTDFSCILVAVCAVITPALRAVPELLIKPRESOVKKMM
MMPGILAILASTFRSNAIDIKTLFGLIIVAITVAHLGRRRLISVAGTIVEV
GLVNLV"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGAGTGTGCGCATCCAGAGAGCCCTG 60
101 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGAGTGTGCGCATCCAGAGAGCCCTG 160
61 GAACCAATGATTAAGATTATCGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 120
161 GAACCAATGATTAAGATTATCGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 220
121 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 180
221 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 280
181 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 240
281 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 340
241 ATGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 300
341 ATGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 400
301 TTGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 360
401 TTGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 460
361 AACCCATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 420
461 AACCCATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 520
421 GGGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 480
521 GGGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 580
481 TCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 540
581 TCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 640
541 ATTAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 600
641 ATTAAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 700
601 CGAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 660
701 CGAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 760
661 GTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 720
761 GTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 820
721 ATCCGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 753
821 ATCCGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 853

RESULT 10
BD014991 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014991
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,

method of isolating the same and utilization thereof.
ACCESSION BD014991
VERSION BD014991.1 GI:2255798
KEYWORDS JP 2001169788-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kernerukunehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
METHOD OF isolating the same and utilization thereof
PATENT: JP 2001169788-A 1 26-JUN-2001;
DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
OS Corynebacterium glutamicum ATCC14752
FN JP 2001169788-A/1
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KERNERUKUNEHITO, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
PFEIFFERLE
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC C12N15/09, C12R1/15, (C12N1/21, C12R1/15), (C12P13/06, C12R1/15),
PC C12N15/00,
PC (C12N15/00, C12R1/15)
CC brnE
CC brnE
FH Key
FT gene
FT Location/Qualifiers
1. .1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGAGTGTGCGCATCCAGAGAGCCCTG 60
101 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGAGTGTGCGCATCCAGAGAGCCCTG 160
61 GAACCAATGATTAAGATTATCGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 120
161 GAACCAATGATTAAGATTATCGCGCTACGAAATCGCGCAAGGTCTAAAACTCCCTT 220
121 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 180
221 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 280
181 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 240
281 GGTGAGAGTTTGGGAGATGTAACCGATGATGAGTTGATGCTCTGATTAATCAATAC 340
241 ATGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 300
341 ATGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 400
301 TTGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 360
401 TTGCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 460
361 AACCCATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 420
461 AACCCATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 520
421 GGGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 480
521 GGGGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 580
481 TCTGATGATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 540

Db 581 TCCCTACTGGGATTCGAGGCTCTACCGAGTGGCGATGCGAGATTGATTCCTTTGAA 640

Qy 541 ATTAAGGCGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGACTTTGATTCCTGC 600

Db 641 ATTAAGGCGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGACTTTGATTCCTGC 700

Qy 601 CGAAGGAAAGACATCCCTCTCTGCTGCTGCGAGGTTTGAGCTTCAACATTGCTCT 660

Db 701 CGAAGGAAAGACATCCCTCTCTGCTGCTGCGAGGTTTGAGCTTCAACATTGCTCT 760

Qy 661 GTGTAATTCAGATCGAGCCCTATTTTGCGGCGCTGCTGATCTTCTTGAGTAC 720

Db 761 GTGTAATTCAGATCGAGCCCTATTTTGCGGCGCTGCTGATCTTCTTGAGTAC 820

Qy 721 ATCCGTAATCTTCTTCTTGAGAAAGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTCTTGAGAAAGCTGCTAA 853

RESULT 11

BD014994 1271 bp DNA linear PAT 27-AUG-2002

LOCUS Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof.

DEFINITION

ACCESSION BD014994.1 GI:22555801

VERSION JP 2001169788-A/4.

KEYWORDS unidentifed

SOURCE unidentifed

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1271)

AUTHORS KernerKnecht,N., Sahn,H., Eggering,L. and Pfefferle,W.

TITLE Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof

JOURNAL DEGRUSA HUBLS AG,FORSCHUNGSZENTRUM JUELICH GMBH

COMMENT OS Corynebacterium glutamicum ATCC13032

PN JP 2001169788-A/4

PD 26-JUN-2001

PF 24-OCT-2000 JP 2000324315

PR 27-OCT-1999 DE 19951708.8

PI NICOLE KERNERKNECHT,HERMANN SAHM,LOTHAR EGGERING,WALTER PI PFEFFERLE

PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/06//

PC (C12N15/09,C12R1:15), (C12N1/21,C12R1:15), (C12P13/06,C12R1:15),

PC C12N15/00,

CC (C12N15/00,C12R1:15)

CC brnF

CC key

PH Key

FT gene

FT Location/Qualifiers

1..1271

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;

Best Local Similarity 100.0%; Pred. No. 1e-184;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCCAAAGGAGCCG 60

Db 101 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCCAAAGGAGCCG 160

Qy 61 GAACCAATGATTAAGGTATTCGGCGCTACGAAATCCGCGAAGGTCTAAAAACCTCCCT 120

Db 161 GAACCAATGATTAAGGTATTCGGCGCTACGAAATCCGCGAAGGTCTAAAAACCTCCCT 220

Qy 121 GCTGAGGTTGGGCAATGACCCGATTCGATTCGTTGGTCTCTTGTTTCAATAC 180

Db 221 GCTGAGGTTGGGCAATGACCCGATTCGATTCGAGGTTGGTCTCTTGTTTCAATAC 280

Qy 181 GGTACGAATGATGGGACACCCCACTGTTTCCGCGCTGATTTTGCGGCGCTCCACCGAA 240

Db 281 GGTACGAATGATGGGACACCCCACTGTTTCCGCGCTGATTTTGCGGCGCTCCACCGAA 340

Qy 241 ATGCTGATACGCGCTCTGTTGTGAGCGAGCGCCCTGCGGCAATCGCGCTACACAA 300

Db 341 ATGCTGATACGCGCTCTGTTGTGAGCGAGCGCCCTGCGGCAATCGCGCTACACAA 400

Qy 301 TTGCTGATACCTTCCGCAAGTATCTATGGTTTCAATCCCGTGAATGGTCAAA 360

Db 401 TTGCTGATACCTTCCGCAAGTATCTATGGTTTCAATCCCGTGAATGGTCAAA 460

Qy 361 AACCCATTGCGCGTTTCTATTCGTTTTCGCGCTATCGAGAAAGCTACGAGTCACT 420

Db 461 AACCCATTGCGCGTTTCTATTCGTTTTCGCGCTATCGAGAAAGCTACGAGTCACT 520

Qy 421 GCGGCGAGCGCGAGCTGCTGCTGCGGCGAGCTTATCTCAATGCAATAGGTTTCA 480

Db 521 GCGGCGAGCGCGAGCTGCTGCTGCGGCGAGCTTATCTCAATGCAATAGGTTTCA 580

Qy 481 TCCCTACTGGGATTCGAGGCTCTACCGAGTGGCGATGCGAGATTGATTCCTTTGAA 540

Db 581 TCCCTACTGGGATTCGAGGCTCTACCGAGTGGCGATGCGAGATTGATTCCTTTGAA 640

Qy 541 ATTAAGGCGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGACTTTGATTCCTGC 600

Db 641 ATTAAGGCGCTCGAGTTCGCCCTTGTCTCTCTTTGTACGCTGACTTTGATTCCTGC 700

Qy 601 CGAAGGAAAGACATCCCTCTCTGCTGCTGCGAGGTTTGAGCTTCAACATTGCTCT 660

Db 701 CGAAGGAAAGACATCCCTCTCTGCTGCTGCGAGGTTTGAGCTTCAACATTGCTCT 760

Qy 661 GTGTAATTCAGATCGAGCCCTATTTTGCGGCGCTGCTGATCTTCTTGAGTAC 720

Db 761 GTGTAATTCAGATCGAGCCCTATTTTGCGGCGCTGCTGATCTTCTTGAGTAC 820

Qy 721 ATCCGTAATCTTCTTCTTGAGAAAGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTCTTGAGAAAGCTGCTAA 853

RESULT 12

AP454053 2105 bp DNA linear BCT 27-JUN-2002

LOCUS Corynebacterium glutamicum BrnF (brnF), BrnF (brnF), and Lrp-1-like

DEFINITION

ACCESSION AP454053

VERSION AP454053.1 GI:21311379

KEYWORDS

SOURCE Corynebacterium glutamicum

ORGANISM Corynebacterium glutamicum

REFERENCE 1 (bases 1 to 2105)

AUTHORS KernerKnecht,N., Sahn,H., Yen,M.R., Patek,M., Sater Jr,M.H. Jr. and Eggering,L.

TITLE Export of L-isoleucine from Corynebacterium glutamicum: a two-gene-encoded member of a new translocator family

JOURNAL J. Bacteriol. 184 (14), 3947-3956 (2002)

PUBMED 12081967

MEDLINE 22077265

JOURNAL 2 (bases 1 to 2105)

REFERENCE KernerKnecht,N., Eggering,L. and Sahn,H.

AUTHORS Direct Submission

TITLE Submitted (07-NOV-2001) IBF-1, Forschungszentrum, Leo-Brandt Str., Juelich 52425, Germany

JOURNAL

FEATURES

source

1..2105

Location/Qualifiers

/organism="Corynebacterium glutamicum"

/mol_type="genomic DNA"

gene	/db_xref="taxon:1719"	
CDS	complement(385..711)	
	/gene="brn"	
	complement(385..711)	
	/gene="brn"	
	/note="LIV-E transporter component 1; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="Brn"	
	/protein_id="AA04685.1"	
	/db_xref="GI:21311380"	
	/translation="MTDPSCLLVNCAVITPALRAVPELILKPLRESQFVKMMWMPAGILALITRSTRNAIDKITPGLINAVITVAHLGRRITLISVAGTIVGLVNL"	
gene	complement(708..1463)	
CDS	/gene="brn"	
	complement(708..1463)	
	/gene="brn"	
	/note="LIV-E transporter component 2; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="Brn"	
	/protein_id="AA04686.1"	
	/db_xref="GI:21311381"	
	/translation="WKTQRIHSLSESPSKALPPDKGYRVEINQGLTSIAGLGVPIGIAFGLVIOGYBWWAPLPSGLIFAGSTEMVILVVGASAPLGLATTLVNFRRVFAFSPLEHVNPIARFVSFALIDEAVAVAPAGSARLLSKGIARHSLVWVGGLGVAIAELIPFELKGLFALCSFVTLTLDSCRTKQIPSLLAGSFTLALVIFGQALFEALILFLILIRYFLKAK"	
gene	1538..1993	
CDS	/gene="lrp"	
	1538..1993	
	/gene="lrp"	
	/note="Lrp"	
	/codon_start=1	
	/transl_table=11	
	/product="Lrp-like regulator"	
	/protein_id="AA04687.1"	
	/db_xref="GI:21311382"	
	/translation="WKLSDIRATIAESANRISNLALADKHLTPPCRRORLEAEGILGSADHLPVWNRGEFVTVLISNDRSTVDNFSVAQHDVEVLEHLRFGSPDYFVRIGVADLEAYEQFLSHIQTVGIAKISRFAMKVVKPARPV"	
ORIGIN		
Query Match	100.0%; Score 753; DB 1; Length 2105;	
Best Local Similarity	100.0%; Pred. No. 1e-184;	
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GTGCAAAAAGCAGAGATTCATTCAGCCTGAGAGTGTGCGCATCCAGGAGACCCCTG 60	
DB	1463 GTGCAAAAAGCAGAGATTCATTCAGCCTGAGAGTGTGCGCATCCAGGAGACCCCTG 1404	
QY	61 GAACCGAGATTAAGTATTCGGGCTGCGAAATCGGACAGGCTCTAAACCTCCCT 120	
DB	1403 GAACCGAGATTAAGTATTCGGGCTGCGAAATCGGACAGGCTCTAAACCTCCCT 1344	
QY	121 GGTGAGGTTTGGGCAATGATCCGATTTGCGTTTGCTCTTGTTATTCATATC 180	
DB	1343 GGTGAGGTTTGGGCAATGATCCGATTTGCGTTTGCTCTTGTTATTCATATC 1284	
QY	181 GGTGAGATGTTGGGCAAGCCCACTGTTTCCGCTCATTTTGGCGGCTCCACCGAA 240	
DB	1283 GGTGAGATGTTGGGCAAGCCCACTGTTTCCGCTCATTTTGGCGGCTCCACCGAA 1224	
QY	241 ATGCTGATCGCCTCGTTGTGGGCGAGCGCCCTGCGGCGCATCGCGTCAACCA 300	
DB	1223 ATGCTGATCGCCTCGTTGTGGGCGAGCGCCCTGCGGCGCATCGCGTCAACCA 1164	
QY	301 TTGCTGGTAACTTCGCGCAGATTTCTATGCTTTTCAATCCGCTGCATGTGTCAAA 360	
DB	1163 TTGCTGGTAACTTCGCGCAGATTTCTATGCTTTTCAATCCGCTGCATGTGTCAAA 1104	
QY	361 AACCCATTGCCGTTTCTATTGCGTTTTCGGGCTTATGAGAGAGCTTACGAGTCACT 420	
DB	1103 AACCCATTGCCGTTTCTATTGCGTTTTCGGGCTTATGAGAGAGCTTACGAGTCACT 1044	
QY	421 GCGGACAGGCGCCGACAGTGTGTCGCGGCGACATTATCTCAATGCAATAGCTTTCAC 480	
DB	1043 GCGGACAGGCGCCGACAGTGTGTCGCGGCGACATTATCTCAATGCAATAGCTTTCAC 984	
QY	481 TCTTACTGGTATTGCGCGGCTCTACCGAGATGCGCATGCGAGATTTCTTTGAA 540	
DB	983 TCTTACTGGTATTGCGCGGCTCTACCGAGATGCGCATGCGAGATTTCTTTGAA 924	
QY	541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTGAATCTGCG 600	
DB	923 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTGAATCTGCG 864	
QY	601 CGAAGAAAAGAGATCCCTTCTCTGCTGCTGCGAGATTGAGCTTCAACATTCCTTT 660	
DB	863 CGAAGAAAAGAGATCCCTTCTCTGCTGCTGCGAGATTGAGCTTCAACATTCCTTT 804	
QY	661 GTGTAATTCAGATCAGAGCCCTATTGCGGCGCTGCTATCTTTGGTCTGTTGAC 720	
DB	803 GTGTAATTCAGATCAGAGCCCTATTGCGGCGCTGCTATCTTTGGTCTGTTGAC 744	
QY	721 ATCCGTAATTCCTTCTTGGAAGAGCTGTAA 753	
DB	743 ATCCGTAATTCCTTCTTGGAAGAGCTGTAA 711	
RESULT 13		
AP005274	340000 bp DNA linear BCT 08-AUG-2002	
LOCUS	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10	
DEFINITION	AP005274 BAO00036	
ACCESSION	AP005274.1 GI:21322764	
VERSION		
KEYWORDS		
SOURCE	Corynebacterium glutamicum ATCC 13032	
ORGANISM	Corynebacterium glutamicum ATCC 13032	
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.	
AUTHORS	Nakagawa, S.	
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 340000)	
AUTHORS	Nakagawa, S.	
TITLE	Direct Submission	
JOURNAL	Submitted (04-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com, Tel: 81-44-829-3831, Fax: 81-44-813-1651)	
COMMENT	This sequence is deposited by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.	
FEATURES	Location/Qualifiers	
source	1..340000	
	/organism="Corynebacterium glutamicum ATCC 13032"	
	/mol_type="genomic DNA"	
	/strain="ATCC 13032"	
	/db_xref="taxon:196627"	
	/note="ATCC 13032"	
gene	1..1575	
	/gene="Cgl10001"	
CDS	1..1575	
	/gene="Cgl10001"	
	/note="PF00308: Bacterial dnaA protein	
	TIGR00362: dnaA: chromosomal replication initiator protein	
	dnaA"	
	/codon_start=1	
	/transl_table=11	
	/product="Atpase involved in DNA replication initiation"	
	/protein_id="BAB97394.1"	

```
/db_xref="GI:21322765"
/translacion="MSQNSSSLLETRWQVADLTLLSQADSGPDLTPROAVNI L
KTAIVDGIATVSTPNAAKNTIENDLGLATRVLSLRGKSFSLAVSEPEQIIPET
PQOEFKYPQDPAVVISNKAPOVEVGEASTDQWERTHSAPAPHPADPE
PELATPQRIPRETPAHNPVREVSINPKYTFEESFVIGPPNRFANAAVAASPAKFN
PLFISGSGIKGTHLHVAAGYAOELQPOLRIKYSSEFTNDYINSVDRDROEFKR
RYNBDILAVDDIOFLAGKEGTQEEFFHFNALHOADKOIILSDRPPKOLITLEDRL
TRVREGGILTDIOPDLETRIILKKAQOTDGTDRYELIISRFSSIRLEEGAL
IKVYASSLINOPIDKEMALVALRDILPEPEMETTAPIVMEVTAVEFESIDTRCA
GKTRAVAHKQAMTICRELTDMSLPKIGDVGSGKHITWTADKIKQEMTERKDTY
DEIQQLTQIKSRGRN"
complement(1594..1920)
/gene="Cg10002"
/complement(1594..1920)
/gene="Cg10002"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translacion="MIVDNFVYLMIDFKVRIILSPYCADVSPAYPOFLGSCAGKFRG
ITPPVTKLELVQNCYATVANKVYDNVNSAFLGSPQNQCTPKLQKQPSDEGSSRAA
EINPVI"
2292..3476
/gene="Cg10003"
2292..3476
/gene="Cg10003"
/EC_number="2.7.7.7"
/ncfe="pf00712:DNA polymerase III beta subunit, N-terminal
domain
pf02767:DNA polymerase III beta subunit, central domain
pf02768:DNA polymerase III beta subunit, C-terminal domain
TIGR00663:dan: DNA polymerase III, beta subunit"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translacion="MESQVSEFVAPEDIVTAVAYARNLPKTPQPIRAMLITADD
EGLEAGYDVYSTRVLSAEOQGRIVAGKLSLSEITGSLPNKRVPRIDGSAFV
TCGSSRFELPLILDYPMPLPKLPAGTGSINAKLFTSAVSQVSAAGXDSLPMLTGV
SMEIVGQINILAAITDFRLALRTFEWANDPELVNGLIIPARTLLDNASLDSGLDS
IDIVAGTGQVGRBEGFVHTDRETTTLADPNNIAPLIPKEHTAISVEIAPLV
DAIRVSIYAEPAACIVIHRSQCVILTGAATAGAAEETLPCATGKELTIAAPGY
LKDGLSVPTSRVAFGFTSPSPAINIIPPEEMPSANENGITQTPDTTYTILMVR L
Pg"
3585..4769
/gene="Cg10004"
3585..4769
/gene="Cg10004"
/ncfe="TIGR00611:recf: Recf protein"
/codon_start=1
/transl_table=11
/product="Recombination DNA repair Atpase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translacion="MHIRSLLELDYRSMPELKYDLEPGITVFIGNGFGKTNIVEAIG
YLAHLSHRVSSDAPLYRAHENAARSAVAVNGRGLAHLIKPRAAQAALNPKY
RTPRELIGVKTITLAFAPEDIALYKGPARERRRLDITITROPMAQVADDPKTY
RNALTKTATLAKRGGTGERGAALSTLTDMDQALRLGAEVAAFPALNLEGPKIT
EAYTTIAPESRPAVANKTITDQSLQSFSPDGIIIEATLVELAKRREIERSGSL
VGPRHDVDMLADOPAKGFSHGETWSPALSLRIAEFNLSGDGTPLLIIIDVSE
LDAGREKIVGIAQVEOVYITAAVDDIPENIKKYLTAQHTVTVDGTGRIISLDV
Op"
4766..5302
/gene="Cg10005"
4766..5302
/gene="Cg10005"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"

```

```
/translacion="MTDPIEOAFERIRAEAMRNGSVPLDNKDAFRPPAPGVEK
RKGRASGLDGRKRVGAEISGLVANKELIRRGKQDIAGWATSNBELVAKIA
QHTRVEMITDKKLPITCDSTAMATINLRMQRQLQVIAKVGPNITTELRITGPQAS
WRKQPLVAKRGPRDITYG"
5354..7489
/gene="Cg10006"
5354..7489
/gene="Cg10006"
/ncfe="pf00204:DNA topoisomerase II (N-terminal region)
pf00986:DNA gyrase B subunit, carboxyl terminus
pf01751:Toprim domain
pf02518:Histidine kinase", DNA gyrase B-, phytochrome-like
Atpase
TIGR01055:pare Gneg: DNA topoisomerase IV, B subunit
TIGR01058:pare Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/transl_table=11
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translacion="MRGTTWGPKRVRMKRLYISSECSLKVANTEHNYDASITILE
GLEAYRRPBGWYIGTGPBRLHILWVYDNSVDAMAGHATKEVTLIDEGVOYVD
DGRGIPVDMHPSGAPTYQVYVMTOLHAGKXDSIYAVSGSLHGVGISVYNALSTRVEA
DIRLHGHWYQNFESVYDYLISGNARAGTITIRWPDALETETEFPEPETSRLQ
EMAFLNKGITITLTNDNATDELELELAEQSTALELIDEIDNETLEVEETDAK
KPKREKKKIIFYHPNGLEDVHYLNRSKNIHISYISFEAKGDHVEVAAWONSXYK
ESVITFANTINTEREGGTHEEGRSALSTLNNRVARBKLIKEXANLITGDCDEGISA
VIVRVDPOFEGQTKLGNTEIKSVQGMANEHIGHMLEANPAKVIINKAVASA
QAPLAAKARDIVRKSAITDGLPGCLACRSDKESLYVEGDSGSAKSGRD
SMRPAIILPLGKILNFKALDYLNKAEVQALITLGCISHPBPINKLRHYKIYM
ADADVQGHATLTLTLFRFMDLVABGVITLAQPLVCLKRQREBPGRAYDERD
EQLNEGIAARCKINKDQIGRYGLGEMNASLEMTTMDPTVILRRVDITDQRADE
LFSILMGDDVVAARSFTRNAKQVFLDI"
7830..8798
/gene="Cg10007"
7830..8798
/gene="Cg10007"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translacion="MIOTYILKIOSLKAQREMLKTLINDLGHVTSFPTTAHP
LPDGSSTPVQIMASDNDSQLVNLMIPFGCGYYPRLAALAKKAGHVAIGLRQGG
QSSAKARSKQMGTHDASVDFPQIAPAKKALDLBEGHMRFLSHMSGQICLPA
RPEAKKYNLARIFGVAGSGSPFTFSPKWKRLGLAVLLGGIGGHI VGPGRVLAK
DLVGYGROSGTHMRERFRKNSLDTLQODINVEVMKVSIIPTFSCPDDECP
QASIDGLASFVPAQIQIEIPEALGHRWARBPESYVCLFLBOAL"
complement(8795..9466)
/gene="Cg10008"
complement(8795..9466)
/gene="Cg10008"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translacion="MSNPNNKPLDPAQRFSLPRWMAFVGFIFMSPAIGITVHAGIG
TTTISLPVYWTASGSLSKWTITYRGNMILCOIYLRQFQPMQVOLLMLFEGF
LCDLSQITTAQDNTFVAMITVYSITLMSISGFYCYAPNTIFIGSGEIVSALYK
FPNVEFGIMQIYDMTIVSVAAIISWTMGGLGVREGTVPAFFIGFQRMKLYL
RSIGH"
complement(9471..9914)
/gene="Cg10009"
complement(9471..9914)
/gene="Cg10009"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"

```

Query Match	100.0%;	Score 753;	DB 1;	Length 340000;
Best Local Similarity	100.0%;	Pred. No. 1.2e-184;		
Matches 753; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

LOCUS	AX120085	349980 bp	DNA	linear	PAT 11-MAY-2001
DEFINITION	Sequence 1 from Patent EP108790.				
ACCESSION	AX120085	AX114121			
VERSION	AX120085.1	GI114036800			
KEYWORDS					
SOURCE					
ORGANISM	Corynebacterium glutamicum				
	Corynebacterium glutamicum				
	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;				
	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.				
REFERENCE	1				
AUTHORS	Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,				
	Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.				
TITLE	Novel polynucleotides				
JOURNAL	Patent: EP 1108790-A 1 20-JUN-2001;				
	KYOWA HAKKO KOGYO CO., LTD. (JP)				
FEATURES	Location/Qualifiers				
source	1..349980				

Query Match	100.0%	Score 753;	DB 6;	Length 349980;
Best Local Similarity	100.0%	Pred. No. 1.2e-184;		
Matches 753; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	GTGCAAAAAACGCAAGATTCATTCAAGCCGTGAGAGGTGGCCATCAAGAGCAGCCCTG	60
Dp	276829	GTGCAAAAAACGCAAGATTCATTCAAGCCGTGAGAGGTGGCCATCAAGAGCAGCCCTG	276888
QY	61	GAACCCGATGATTAAGGTTATCGGCGCTACGAAATCGCGCAAGTCTAAAACCTCCCTT	120
Dp	276889	GAACCCGATGATTAAGGTTATCGGCGCTACGAAATCGCGCAAGTCTAAAACCTCCCTT	276948
QY	121	GCAGCAGGTTGGGGGCAATGACCGAATGGGTAATGGGTTGGTCTCTGGTTATTCAATAC	180
Dp	276949	GCAGCAGGTTGGGGGCAATGACCGAATGGGTAATGGGTTGGTCTCTGGTTATTCAATAC	277008
QY	181	GGCTACGAATGATGAGGCAAGCCCACTGTTTTCGGCTCATATTTTCGGGAGTCCACCGAA	240
Dp	277009	GGCTACGAATGATGAGGCAAGCCCACTGTTTTCGGCTCATATTTTCGGGAGTCCACCGAA	277068
QY	241	ATGCTGATCATGCGCCCTCGTTGTGGGCGCAGCGCCCTTGCGCGCATGCGGCTCACACA	300
Dp	277069	ATGCTGATCATGCGCCCTCGTTGTGGGCGCAGCGCCCTTGCGCGCATGCGGCTCACACA	277128
QY	301	TTCCTGGTGAACCTTCGCGCAGATTCCTATGCGTTTTCATTCGCGCTGCATGAGGTCAAA	360
Dp	277129	TTCCTGGTGAACCTTCGCGCAGATTCCTATGCGTTTTCATTCGCGCTGCATGAGGTCAAA	277188
QY	361	AACCCCATGCGCCGTTTCTATTGCGTTTTCGCGCTTATCGAAGAGCTTACGCACTACT	420
Dp	277189	AACCCCATGCGCCGTTTCTATTGCGTTTTCGCGCTTATCGAAGAGCTTACGCACTACT	277248
QY	421	GCGGCGAGGCGCCGAGGCTGTGCGCGTGGGCACTTATCTCAATGCAATATGCGTTTAC	480
Dp	277249	GCGGCGAGGCGCCGAGGCTGTGCGCGTGGGCACTTATCTCAATGCAATATGCGTTTAC	277308
QY	481	TCCCATGAGGATATCGGCGAGTCTCAACCGGAGTGGCGATCGCAGATGATTCCTTTGAA	540
Dp	277309	TCCCATGAGGATATCGGCGAGTCTCAACCGGAGTGGCGATCGCAGATGATTCCTTTGAA	277368
QY	541	ATTAAAGGCGCTCGAGTTGCGCCCTTGCCTCTCTTTGTACGCTGACTTTGGAATTCCTGC	600
Dp	277369	ATTAAAGGCGCTCGAGTTGCGCCCTTGCCTCTCTTTGTACGCTGACTTTGGAATTCCTGC	277428
QY	601	CGAAGCAAAAAAGACGATCCCTTCTGCTGCTCTCGAGAGTTTGAAGTTACATATTCCTT	660
Dp	277429	CGAAGCAAAAAAGACGATCCCTTCTGCTGCTCTCGAGAGTTTGAAGTTACATATTCCTT	277488
QY	661	GTAGTATTTCAAGTCAAGGCCCTATTATGGCGCGCTGATCTCTTTGGAGTCTTGGACC	720
Dp	277489	GTAGTATTTCAAGTCAAGGCCCTATTATGGCGCGCTGATCTCTTTGGAGTCTTGGACC	277548
QY	721	ATCCGGAATCTTCTTCTTGGAAAAGCTGCTAAA	753
Dp	277549	ATCCGGAATCTTCTTCTTGGAAAAGCTGCTAAA	277581

QY	1	GTCTCAAAAAACGCAAGAATCATCTCAAGCCGTGAGAGTGTGGCATTCAGAGCAGCCCTG	60
Dd	276829	GTCTCAAAAAACGCAAGAATCATCTCAAGCCGTGAGAGTGTGGCATTCAGAGCAGCCCTG	276888
QY	61	GAACCCAGATGATTAAGTTATTCGGGGCTTACGAATTCGGCGCAAGTGTAAAACTCCCTT	120
Dd	276889	GAACCCAGATGATTAAGTTATTCGGGGCTTACGAATTCGGCGCAAGTGTAAAACTCCCTT	276948
QY	121	GCTGCAAGTTTGGGCAATGTAACCCGATTTGGTATTTGGCTTTGGTATTCAATAC	180
Dd	276949	GCTGCAAGTTTGGGCAATGTAACCCGATTTGGTATTTGGCTTTGGTATTCAATAC	277008
QY	181	GGCTACGAATGATGAGGGCAGCCCCACGTGTTTCGGGCTCGATTTTCGGGGGCTCCACCGAA	240
Dd	277009	GGCTACGAATGATGAGGGCAGCCCCACGTGTTTCGGGCTCGATTTTCGGGGGCTCCACCGAA	277068
QY	241	ATGCTGTCAATGCGCCCTGTGTGGGGCGAGCGCCCTTGAGGCGCATGCGCTCACCA	300
Dd	277069	ATGCTGTCAATGCGCCCTGTGTGGGGCGAGCGCCCTTGAGGCGCATGCGCTCACCA	277128
QY	301	TTGCTGGTGAATCTTCGCGCAGCATATCTATAGCGTTTCAATTCGCGCTCATGTGGTCAAA	360
Dd	277129	TTGCTGGTGAATCTTCGCGCAGCATATCTATAGCGTTTCAATTCGCGCTCATGTGGTCAAA	277188
QY	361	AAACCCATTGCCCCGTTTCTATTGCGTTTTCGCGCTTATACGAGAGGCTACGAGTCACT	420
Dd	277189	AAACCCATTGCCCCGTTTCTATTGCGTTTTCGCGCTTATACGAGAGGCTACGAGTCACT	277248
QY	421	GGCGCCAGGCGCCGACAGCTGTGTGGCGTGGGCACTATCTCAATGCAATAGCGTTTCA	480
Dd	277249	GGCGCCAGGCGCCGACAGCTGTGTGGCGTGGGCACTATCTCAATGCAATAGCGTTTCA	277308
QY	481	TTCCTACTGGGTATTTCGGCGGCTCACCCGAGAGGGGAGATCGAGAGATGATTCCTTTGAA	540
Dd	277309	TTCCTACTGGGTATTTCGGCGGCTCACCCGAGAGGGGAGATCGAGAGATGATTCCTTTGAA	277368
QY	541	ATTAAAGGCTCGAGTTTGGCCCTTGTCTCTCTTTGTACGCTGACTTTGAGATTCTGC	600
Dd	277369	ATTAAAGGCTCGAGTTTGGCCCTTGTCTCTCTTTGTACGCTGACTTTGAGATTCTGC	277428
QY	601	CGAACGAAAAAGAGATCCCTTCTGTGCTGTCTCGAGAGTTTAAAGCTTACCACTTGGCTTT	660
Dd	277429	CGAACGAAAAAGAGATCCCTTCTGTGCTGTCTCGAGAGTTTAAAGCTTACCACTTGGCTTT	277488
QY	661	GTGGTAATTCCAGGTCAAGCCCTATTTCGGCGCGCTGTGATTTCTTGGGTCTGTGACC	720

CC branched-chain L-aa by fermentation of coryneform bacteria in which the
CC bms and/or bms genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the bms and/or
CC bms genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 14752 bms protein described in the
CC method of the invention

XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-223;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGCATCCAGGACGCTCG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGCATCCAGGACGCTCG 60
QY 61 GAACCAAGATGATTAAGGTTATCGCGCTTACGAATCGCGCAAGGTCTAAACCTCCCT 120
DB 61 GAACCAAGATGATTAAGGTTATCGCGCTTACGAATCGCGCAAGGTCTAAACCTCCCT 120
QY 121 GCTGCAAGTTGGGCGCATGTAACCGATGATGATGCTGTTGCTGTTATTCATAC 180
DB 121 GCTGCAAGTTGGGCGCATGTAACCGATGATGATGCTGTTGCTGTTATTCATAC 180
QY 181 GGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 AACCCATGCGCGCTTCTATTCGCTTTCGCTTATCGACGACGCTTACGACGCT 420
DB 361 AACCCATGCGCGCTTCTATTCGCTTTCGCTTATCGACGACGCTTACGACGCT 420
QY 421 GCGGCGAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 GCGGCGAGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 TCCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 TCCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 ATTAAAGGCTTCGAGTTCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 600
DB 541 ATTAAAGGCTTCGAGTTCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 600
QY 601 CGAACGAAAAAGCATGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 CGAACGAAAAAGCATGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GTGCTAATTCAGGTCAGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 720
DB 661 GTGCTAATTCAGGTCAGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 720
QY 721 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
DB 721 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753

RESULT 2
AAH65254

ID AAH65254 standard; DNA; 753 BP.
AC AAH65254;
XX 26-SEP-2001 (first entry)
DE C glutamicum coding sequence fragment SEQ ID NO: 289.
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; ds.
XX Corynebacterium glutamicum.
XX EPI108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99UP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX P-PSDB; AAG90035.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX Claim 8; SEQ ID NO 289; 246pp + Sequence listing; English.
XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Coryneform bacterium, and identifying a homologue of a gene derived from
XX Coryneform bacterium. Coryneform bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office

Seq Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 5; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-223;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGCATCCAGGACGCTCG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGCATCCAGGACGCTCG 60
QY 61 GAACCAAGATGATTAAGGTTATCGCGCTTACGAATCGCGCAAGGTCTAAACCTCCCT 120
DB 61 GAACCAAGATGATTAAGGTTATCGCGCTTACGAATCGCGCAAGGTCTAAACCTCCCT 120
QY 121 GCTGCAAGTTGGGCGCATGTAACCGATGATGATGATGATGATGATGATGATGAT 180
DB 121 GCTGCAAGTTGGGCGCATGTAACCGATGATGATGATGATGATGATGATGATGAT 180
QY 181 GGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 GGCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300

Db 241 ATGCTGATCATCGCCCTGCTGTGTGGGCGCAGCCGCCGCGCATCGGCTCACACACA 300
 Qy 301 TTGCTGATGAACTTCCGCGCAGATATTCATGCGTTTTCATTCGCGCGCATGTCGCAAA 360
 Db 301 TTGCTGATGAACTTCCGCGCAGATATTCATGCGTTTTCATTCGCGCGCATGTCGCAAA 360
 Qy 361 AACCCATTCGCGCGTTTCTATTCGATTCGCGTTTTCATTCGCGCGCATGTCGCAAA 420
 Db 361 AACCCATTCGCGCGTTTCTATTCGATTCGCGTTTTCATTCGCGCGCATGTCGCAAA 420
 Qy 421 GCGGCGCAGGCG 480
 Db 421 GCGGCGCAGGCG 480
 Qy 481 TCCTACTGATGATTCG 540
 Db 481 TCCTACTGATGATTCG 540
 Qy 541 ATTAAGGCGCTCGAGTTCG 600
 Db 541 ATTAAGGCGCTCGAGTTCG 600
 Qy 601 CGAAGGCG 660
 Db 601 CGAAGGCG 660
 Qy 661 GTGCTAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 720
 Db 661 GTGCTAATTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCG 720
 Qy 721 ATCCGCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 753
 Db 721 ATCCGCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 753

RESULT 3

AAH21109 standard; DNA; 1271 BP.

AAH21109;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnB.

L-amino acid production; brnF, brnB; branched-chain amino acid;

coryneform bacterium; leucine; isoleucine; valine; medicine;

animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers

FT CDS 101..856

FT CDS /*tag= a

FT CDS /product= "brnF"

FT CDS 853..1179

FT CDS /*tag= b

FT CDS /product= "brnB"

PN EPI096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000BP-00122057.

PR 27-OCT-1999; 99DE-01051708.

PA (DEGS) DEGUSA AG.

PA (KER) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Kennerknecht N, Eggeling L, Sahm H, Pfeiferle W;

XX

DR WPI; 2001-391595/42.
 XX P-PSDB; AAB86247, AAB86248.
 PT New expert genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 4 (i); Page 13; 23pp; German.
 CC This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (I), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnB and/or brnF genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnB and/or
 CC brnF genes. (I) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (I) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (I)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brnB and brnB proteins described in
 CC the method of the invention
 XX
 SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred No. 2.5e-223;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAACGCAAGATATTCATTCAGCTGAGAGTGTGCGCATCAAGGAGCGCTTG 60
 Db 101 GTGCAAAAACGCAAGATATTCATTCAGCTGAGAGTGTGCGCATCAAGGAGCGCTTG 160
 Qy 61 GAACCAAGATATTAAGGATATTCAGCTGAGAGTGTGCGCATCAAGGAGCGCTTG 120
 Db 161 GAACCAAGATATTAAGGATATTCAGCTGAGAGTGTGCGCATCAAGGAGCGCTTG 220
 Qy 121 GCTGAGGTTGGGCGATGACCCGATGATGATGATGATGATGATGATGATGATGATGATG 180
 Db 221 GCTGAGGTTGGGCGATGACCCGATGATGATGATGATGATGATGATGATGATGATGATG 280
 Qy 181 GGTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 Db 281 GGTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 340
 Qy 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 Db 341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400
 Qy 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
 Db 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 460
 Qy 461 AACCCATTCGCGCGTTTCTATTCGATTCGCGTTTCTATTCGATTCGCGCGTTTCTATTC 420
 Db 461 AACCCATTCGCGCGTTTCTATTCGATTCGCGTTTCTATTCGATTCGCGCGTTTCTATTC 520
 Qy 421 GCGGCGCAGGCG 480
 Db 521 GCGGCGCAGGCG 580
 Qy 481 TCCTACTGATGATTCG 540
 Db 581 TCCTACTGATGATTCG 640
 Qy 541 ATTAAGGCGCTCGAGTTCG 600

Db 641 ATTAAGGCGCTCGAGTTGGCCCTTTGCTCTCTTTGTCACGCTGACTTTGGATTCCGTC 700
 Qy 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 760
 Qy 661 GTGGTAATTCAGGTCAGGCGCTATTTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 720
 Db 761 GTGGTAATTCAGGTCAGGCGCTATTTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 820
 Qy 721 ATCCGTAATCTTCTTCTTGGGAAAAGCTGCTAA 753
 Db 821 ATCCGTAATCTTCTTCTTGGGAAAAGCTGCTAA 853

RESULT 4

AAH21112
ID AAH21112 standard; DNA, 1271 BP.

AAH21112;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnE.

L-amino acid production; brnF; brnE; branched-chain amino acid;
 coryneform bacterium; leucine; isoleucine; valine; medicine;
 animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers
 FH 101..856
 FT /*tag= a
 FT /product= "brnF"
 FT 853..1179
 FT /*tag= b
 FT /product= "brnE"

EP1096010-A1.

02-MAY-2001.

11-OCT-2000; 2000EP-00122057.

27-OCT-1999; 99DE-01051708.

(DEGS) DEGUSA AG.

(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

Kemernecht N, Eggeling L, Sahn H, Pfeifferle W;

WPI; 2001-391595/42.

New export genes from coryneform bacteria, useful for increasing

fermentative production of branched-chain amino acids.

Claim 4 (i); Page 17-18; 23pp; German.

This invention describes a novel isolated polynucleotide (i) containing
 at least one sequence that (i) is 70% identical with a sequence that
 encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 15 consecutive bases from (i)-(iii). The invention also describes (a)
 protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 coryneform microorganisms, especially Corynebacterium, transformed with
 one or more (i), where these are replicative DNA; (c) production of
 branched-chain L-aa by fermentation of coryneform bacteria in which the
 brnF and/or brnE genes (or equivalent sequences) are amplified,
 especially overexpressed; and (d) method for isolating the brnF and/or
 brnE genes. (i) is used for transformation of coryneform bacteria being
 used for fermentative production of branched-chain amino acids,

CC Specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (i) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 13032 brnF and brnE proteins described in
 CC the method of the invention

Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query March 100.0%; Score 753; DB 4; Length 1271;

Best Local Similarity 100.0%; Pred. No. 2.5e-223; Mismatches 0; Gaps 0;

Matches 753; Conservative 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGAGTGTGCGATCCAGGCGAGCCCTG 60
 Db 101 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGAGTGTGCGATCCAGGCGAGCCCTG 160
 Qy 61 GAACCAATGATTAAGATTATGCGCGCTAGAAATGCGCAAGTCTAAACCTCCCTT 120
 Db 161 GAACCAATGATTAAGATTATGCGCGCTAGAAATGCGCAAGTCTAAACCTCCCTT 220
 Qy 121 GCTGCAAGTTGGGAGATGACCGGATTTGATTTGCTGCTGTTATTCATAC 180
 Db 221 GCTGCAAGTTGGGAGATGACCGGATTTGATTTGCTGCTGTTATTCATAC 280
 Qy 181 GGCTAGCAATGTGGGAGAGCCCACTGTTTTCGCGCTGATTTTCGCGGCTCCACCGA 240
 Db 281 GGCTAGCAATGTGGGAGAGCCCACTGTTTTCGCGCTGATTTTCGCGGCTCCACCGA 340
 Qy 241 ATGCTGATCATCGCCCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGCTACACCA 300
 Db 341 ATGCTGATCATCGCCCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGCTACACCA 400
 Qy 301 TTGCTGATGATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGCAAA 360
 Db 401 TTGCTGATGATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGCAAA 460
 Qy 361 AACCCATTCGCGGTTTCAATTCGCTTTTCGCGCTATGAGAGGAGCTACGAGCACT 420
 Db 461 AACCCATTCGCGGTTTCAATTCGCTTTTCGCGCTATGAGAGGAGCTACGAGCACT 520
 Qy 421 GCGGCGAGGCGCGAGGCTGTGCGCGTGGCGAGCTTATCTCAATGCAAAATAGCGTTTCA 480
 Db 521 GCGGCGAGGCGCGAGGCTGTGCGCGTGGCGAGCTTATCTCAATGCAAAATAGCGTTTCA 580
 Qy 481 TCCACTGAGTATTCGCGCGTGTCAACCGAGTGGCGAGTGAATTCCTTTGAA 540
 Db 581 TCCACTGAGTATTCGCGCGTGTCAACCGAGTGGCGAGTGAATTCCTTTGAA 640
 Qy 541 ATTAAGGCGCTCGAGTTGCGCTTGTCTCTCTTTGTCAAGCTGACCTTGAATCTTGC 600
 Db 641 ATTAAGGCGCTCGAGTTGCGCTTGTCTCTCTTTGTCAAGCTGACCTTGAATCTTGC 700
 Qy 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 760
 Qy 661 GTGGTAATTCAGGTCAGGCGCTATTTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 720
 Db 761 GTGGTAATTCAGGTCAGGCGCTATTTTGCGGCGTGTGATCTTCTTGGTCTGTGACC 820
 Qy 721 ATCCGTAATCTTCTTCTTGGGAAAAGCTGCTAA 753
 Db 821 ATCCGTAATCTTCTTCTTGGGAAAAGCTGCTAA 853

RESULT 5

AAH64966
ID AAH64966 standard; DNA, 349980 BP.

AAH64966;

26-SEP-2001 (first entry)

XX DE C glutamicum coding sequence fragment SEQ ID NO: 1.
 XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 XX KM organic acid synthesis; ds.
 XX OS Corynebacterium glutamicum.
 XX PN EP1108790-A2.
 XX PD 20-JUN-2001.
 XX PF 18-DEC-2000; 2000EP-00127688.
 XX PR 16-DEC-1999; 99JP-00377484.
 XX PR 07-APR-2000; 2000JP-00159162.
 XX PR 03-AUG-2000; 2000JP-00280988.
 XX PA (KJOM) KYOMA HAKKO KOGYO KK.
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochiai K, Yokoi H;
 XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX DR WPI; 2001-376931/40.
 XX PS
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 7; SEQ ID NO 1; 246bp + sequence listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of Corynebacterium bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
 CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;
 Query Match 100.0%; Score 753; DB 5; Length 349980;
 Best Local Similarity 100.0%; Fred. No. 3.1e-222;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 AACCCATGCCCCGTTTCTATTGCGTTTTCGGCTTATGAGAAAGCTACGAGTCACT 420
 DB 277189 AACCCATGCCCCGTTTCTATTGCGTTTTCGGCTTATGAGAAAGCTACGAGTCACT 277248
 QY 421 GCGGCCAGGCGCCGACGCTGCTGCGCTGCGGCACTTCTCAATGCAATTAAGCTTTCAC 480
 DB 277249 GCGGCCAGGCGCCGACGCTGCTGCGCTGCGGCACTTCTCAATGCAATTAAGCTTTCAC 277308
 QY 481 TCCATCTGGGTTATTCGGCGGCTCACCGGAGTGGCATGCGAGATTGATTCCTTTGAA 540
 DB 277309 TCCATCTGGGTTATTCGGCGGCTCACCGGAGTGGCATGCGAGATTGATTCCTTTGAA 277368
 QY 541 ATTAAGGCGCTCGAGTTCCGCTTCTGCTCTCTTTGTCAAGCTTTCGATTTCTGCG 600
 DB 277369 ATTAAGGCGCTCGAGTTCCGCTTCTGCTCTCTTTGTCAAGCTTTCGATTTCTGCG 277428
 QY 601 CGAACGAAAGAGAGATCCCTTCTGCTGCTGCTGCGAGTTTGAAGCTTCAACATTGCTCTT 660
 DB 277429 CGAACGAAAGAGAGATCCCTTCTGCTGCTGCTGCGAGTTTGAAGCTTCAACATTGCTCTT 277488
 QY 661 GTGGTAATTCAGAGTCAGGCTTATTTGCGGCGCTGCTGATCTTCTTGGTCTGTTGACC 720
 DB 277489 GTGGTAATTCAGAGTCAGGCTTATTTGCGGCGCTGCTGATCTTCTTGGTCTGTTGACC 277548
 QY 721 ATCCGTAATTCAGAGTCAGGCTTATTTGCGGCGCTGCTGATCTTCTTGGTCTGTTGACC 753
 DB 277549 ATCCGTAATTCAGAGTCAGGCTTATTTGCGGCGCTGCTGATCTTCTTGGTCTGTTGACC 277581

RESULT 6
 ID ACA01969 standard; DNA; 564 BP.
 AC ACA01969;
 DT 04-JUN-2003 (first entry)
 DE C. glutamicum derived ORF SEQ ID 1960.
 KM Coryneform; nucleic acid array; fermentation; culture; ds.
 OS Corynebacterium glutamicum.
 FN DE10128510-A1.
 PD 19-DEC-2002.
 PF 13-JUN-2001; 2001DE-01028510.
 PR 13-JUN-2001; 2001DE-01028510.
 PA (DEGS) DEGUSSA AG.
 PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Huthmacher K;
 DR WPI; 2003-279970/28.
 PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 PS Claim 1; Page 647; 709pp; German.
 CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;

Query Match 73.4%; Score 553; DB 7; Length 564;
Best Local Similarity 99.8%; Pred. No. 2.7e-161;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 189 ATGGTGGAGAGCCCACTGTTTCGGGCTGATTTTCGGGGCTCCACGGAATGCTGAT 248
DB 1 ATGGTGGAGAGCCCACTGTTTCGGGCTGATTTTCGGGGCTCCACGGAATGCTGAT 60
QY 249 CATGCGCTCGTGTGGGGCGCAGCGCCCTGGGCGGCATCGGCTCACCACATTTGCTGAT 308
DB 61 CATGCGCTCGTGTGGGGCGCA-CGCCCTGGGCGGCATCGGCTCACCACATTTGCTGAT 119
QY 309 GAACCTCCCGCAGCTATTTCTAGACCTTTTCATTTCCCGCGCATGTGTGTAATAACCCAT 368
DB 120 GAACCTCCCGCAGCTATTTCTAGACCTTTTCATTTCCCGCGCATGTGTGTAATAACCCAT 179
QY 369 TGCCCGTTTCTATTCGATTTTCGGCTTATCGACGAGCTACGAGTCACTGCGGCGAG 428
DB 180 TGCCCGTTTCTATTCGATTTTCGGCTTATCGACGAGCTACGAGTCACTGCGGCGAG 239
QY 429 GCCCGCAGGCTGATGCGGCTGCGGCTTATCTCAATGCAATAGGCTTCACTCTACTG 488
DB 240 GCCCGCAGGCTGATGCGGCTGCGGCTTATCTCAATGCAATAGGCTTCACTCTACTG 299
QY 489 GGTATTGCGGCTGCTACCGGAGTGGCGATCGCAGAGTGTATCTTTGAAATTAAAGG 548
DB 300 GGTATTGCGGCTGCTACCGGAGTGGCGATCGCAGAGTGTATCTTTGAAATTAAAGG 359
QY 549 CCTGAGTTGCGGCTTGTCTCTCTTTGTACGCTGACTTTGATTTCTGCGGAAGAA 608
DB 360 CCTGAGTTGCGGCTTGTCTCTCTTTGTACGCTGACTTTGATTTCTGCGGAAGAA 419
QY 609 AAGGACGATGCTTCTGCTGCTGCTGCGAGGTTGAGTTCACATGCTTGTGTAT 668
DB 420 AAGGACGATGCTTCTGCTGCTGCTGCGAGTTCGAGTTGAGTTCACATGCTTGTGTAT 479
QY 669 TCAGGTCAGGCTTATTTTGGGCGCTGCTGATTTCTTGGATCTGTGACATCGGTA 728
DB 480 TCAGGTCAGGCTTATTTTGGGCGCTGCTGATTTCTTGGATCTGTGACATCGGTA 539
QY 729 CTCTCTTTGGGAAAGGCTGCTAA 753
DB 540 CTCTCTTTGGGAAAGGCTGCTAA 564

```

RESULT 7

AAf61693/c
ID AAF61693 standard; DNA; 397 BP.

XX AAF61693;

XX 12-JUL-2001 (first entry)

XX C. glutamicum lrp encoding DNA fragment SEQ ID 7.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
XX medicine; animal feed supplement; ds.

XX Corynebacterium glutamicum.

XX EP1090993-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121159.

XX 05-OCT-1999; 99DE-01047792.

XX (DEGSA) DEGUSA-HUELS AG.

PI Moeckel B, Pfeifferle W, Puhler A, Kalinowski J, Bathe B;
XX WPI; 2001-29927/31.

XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.

XX Example 3; Page 16; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention

SEQ Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

Query Match 36.3%; Score 273.6; DB 4; Length 397;
Best Local Similarity 98.6%; Pred. No. 1.7e-74;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAGCAGAGATTCAAGCTGAGGCTGCGCATCAAGGAGCCCTG 60
DB 284 GTGCAAAAAGCAGAGATTCAAGCTGAGGCTGCGCATCAAGGAGCCCTG 225
QY 61 GAACGATGATTAAGTTATGCGGCTACGAAATCGCGCAGGCTTAAAACTCCCTT 120
DB 224 GAACGATGATTAAGTTATGCGGCTACGAAATCGCGCAGGCTTAAAACTCCCTT 165
QY 121 GCTGAGGTTGAGGATGATGACCGGATGATTTGCTGTTGCTCTGTTATTAATAC 180
DB 164 GCTGAGGTTGAGGATGATGACCGGATGATTTGCTGTTGCTCTGTTATTAATAC 105
QY 181 GCTACGATGATGAGGAGCAGCCCACTGTTTCGAGCTGATTTTCGCGGCTCCACGAA 240
DB 104 GCTACGATGATGAGGAGCAGCCCACTGTTTCGAGCTGATTTTCGCGGCTCCACGAA 45
QY 241 ATGCTGATCATGCGGCTGTTGTTGGGCGCAGCGGCGGCTG 280
DB 44 ATGCTGATCATGCGGCTGTTGTTGGGCGCAGCGGCGGCGG 5

```

RESULT 8

AAf61695/c
ID AAF61695 standard; DNA; 778 BP.

XX AAF61695;

XX 12-JUL-2001 (first entry)

XX C. glutamicum lrp encoding DNA fragment SEQ ID 9.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
XX medicine; animal feed supplement; ds.

XX Corynebacterium glutamicum.

XX EP1090993-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121159.

XX 05-OCT-1999; 99DE-01047792.

XX (DEGSA) DEGUSA-HUELS AG

XX Moeckel B, Pfefferle W, Puhler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX

XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX

XX Example 3; Page 17; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention

XX Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;

Query Match 36.3%; Score 273.6; DB 4; Length 778;
Best Local Similarity 98.6%; Pred. No. 2.3e-74;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGCCCTGAGAGTGTCCATCCAGACGACCCCTG 60
DB 284 GTGCAAAAACGCAAGATTCATTCAGCCCTGAGAGTGTCCATCCAGACGACCCCTG 225
QY 61 GAACCATGATGATTAAGATTATGCGGCTACGAAATCGCGAGAGTCTAAAACTTCCTT 120
DB 224 GAACCATGATGATTAAGATTATGCGGCTACGAAATCGCGAGAGTCTAAAACTTCCTT 165
QY 121 GGTGCAAGTTGGGCAATGACCGATTTGATTTGGTTGCTCTGTTGTTATCAATAC 180
DB 164 GGTGCAAGTTGGGCAATGACCGATTTGATTTGGTTGCTCTGTTGTTATCAATAC 105
QY 181 GGCTACGAATGCTGGGCAAGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 240
DB 104 GGCTACGAATGCTGGGCAAGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 45
QY 241 ATGCTGCTCATCGCCCTCTGTTTGGGGCGACGCGCCCTG 280
DB 44 ATGCTGCTCATCGCCCTCTGTTTGGGGCGACGCGCCCTG 5

RESULT 9

AAH68373
ID AAH68373 standard; DNA; 711 BP.

XX AAH68373;
AC
DT 26-SEP-2001 (first entry)
XX

DE C glutamicum coding sequence fragment SEQ ID NO: 3408.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX

XX Corynebacterium glutamicum.
OS
XX

PN EP108790-A2.
XX

PD 20-JUN-2001.
XX

XX 18-DEC-2000; 2000EP-00127688.
PF

XX 16-DEC-1999; 99JP-00377484.
PR

07-APR-2000; 2000JP-00159162.
PR

PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX

XX WPI; 2001-376931/40.
XX P-PSDB; JAG93154.
XX

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 3408; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of coryneform bacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Coryneform bacterium, and identifying a homologue of a gene derived from
XX Coryneform bacterium. Coryneform bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office

XX Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 5; Length 711;
Best Local Similarity 51.5%; Pred. No. 6.3e-29;
Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGCAAGTTTGGGCAATGATCCGATTTGATTTGGTTGCTCTGTTGTTATCAATAC 181
DB 53 CGGTGCGTTTGGGCAATGATCCGATTTGATTTGGTTGCTCTGTTGTTATCAATAC 112
QY 182 GCTACGAATGCTGGGCAAGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 241
DB 113 GTTTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 172
QY 242 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
DB 173 TTCTGCAATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 232
QY 302 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
DB 233 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 292
QY 362 ACCCATTTGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
DB 293 CCGGCGCGGCGCGGCGCTATTCACCTACGCGCTTACCGACGACCTACGCGCT 352
QY 422 CGGCGCGGCGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 478
DB 353 CAGCGCGGCGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 412
QY 479 ACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
DB 413 AAGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 472
QY 539 AATTTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 598
DB 473 ATCTAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 532
QY 599 GCGCAAGCAAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 658
DB 533 TCAAAAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 592
QY 659 TTGTGATTTCCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 718

Db 233 TCATGGTGAATTTCCGCCACATTTTCTACGGTCTCACCTTCCGACGCCACCGCATCAAGT 297

the expression profile or expression pattern of a gene derived from mutant or *Corynebacterium*, measuring expression amount and

Corynebacterium, and identifying a homologue of a gene derived from

CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 5; Length 309400;

Best Local Similarity 51.5%; Pred. No. 9.6e-28;

Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGACAGTTTGGGACATGACCCGATTTGATTCGTTGGTCTCTTGGTTATTCATACG 181
 DB 289969 CGGTCCGTTTGGGGCGATTCCTCGGTTTGGCCCTTGGCGTTGATGATCCAGACG 289910
 QY 182 GCTACGAATGATGGGACGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGA 241
 DB 289909 GTTCCGCTGTGTGTGAGCCGATTTCTCTTCTGATCTATGCGGTTGATGAAAT 289850
 QY 242 TGTGTGATCGCCCTGTGTGTGTGGGCGACGCCCTCTGGGCGCATCGCTTACCAAT 301
 DB 289849 TTTCTGGAATCGGCATGTGTACCGCAGTATCGGCCCGTTTTCGGCGGTCGCTGAT 289790
 QY 302 TGTGTGGAATCCGCGCAGTATTCATGCGTTTCAATCCCGCTGCAATGTGTCAAA 361
 DB 289789 TCAATGTAATTTCCGCGCAATTTTCAAGCTCTCACCTTCCACGCCACCGCATGAT 289730
 QY 362 ACCCATTTGCCGTTTCTATTCGTTTCCGCTTATCGAAGGCTTACGAGTCACTG 421
 DB 289729 CCGGCGCGCGCGCGCTATTCACCTAGCGCTTACCAAGAGTCTTACGCGATGTGT 289670
 QY 422 CGGCGAGCGCGGAGGTGTGTGGCGG---GGGACTTATCTCAATGCAAAATAGCTTTC 478
 DB 289669 CACCCCGCCACCTGGGATATCAGTGACGCGGCGCTTACGTTCAAATTTGTGCT 289610
 QY 479 ACTCTACTAGGTATTCGCGGTCTCACCGAGTGGCATTCGAGAGTTGATTCCTTTG 538
 DB 289609 AAGCTGTGTGTTATCCAGGAATATATGGCGCTTGTGTGTAAGTGTGCGCGATG 289550
 QY 539 AAATTAAGGCGCTGAGTTGCGCTTCTCTCTTGTTCAGCGTCACTTGTGATTCCT 598
 DB 289549 ATCTAAAGGCAATGATTTTGCCTGACCGCGCTGTTGTGTGCTGCGGTGGAGCAT 289490
 QY 599 GCCGAACGAAGAAAGCAGATCCCTTCTGTGCTGCAAGTTTGAAGTTACACATTGCTC 658
 DB 289489 TCAAAAATTAAGAGATTAATCGCTCATTAATCCGATGATATGGCTGTGTTCCG 289430
 QY 659 TTGTGTATTCAGGTGAGGCGCTTATTTGGCGCGCTGTGATCTTGTGGTCTGTGA 718
 DB 289429 GTTTTGTGGCGCGCGAGCAATGCTGTGTTATCGTTTGAACAAGTATTTTGTATCTTC 289370
 QY 719 CCATCGGACTTCTTC 735
 DB 289369 TTCTCGCGCTCGCTTC 289353

RESULT 12

ACA01968 standard; DNA; 177 BP.

XX ACA01968;

XX 04-JUN-2003 (first entry)

XX C. glutamicum derived ORF SEQ ID 1959.

XX Coryneform; nucleic acid array; fermentation; culture; ds.

XX Corynebacterium glutamicum.

XX DB10128510-A1.

XX 19-DEC-2002.
 XX 13-JUN-2001; 2001DE-01028510.
 XX 13-JUN-2001; 2001DE-01028510.
 XX (DEGS) DEGUSA AG.

XX Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;

XX WPI; 2003-279970/28.

XX New nucleic acid array useful for monitoring mRNA expression of

XX Corynebacterium glutamicum during fermentation, comprising nucleic acid

XX from Corynebacterium glutamicum.

XX Claim 1; Page 647; 709pp; German.

XX This invention describes a novel nucleic acid array involving

XX Corynebacterium glutamicum polynucleotides. The arrays are used to

XX analyse C. glutamicum, particularly for monitoring a fermentation process

XX to determine expression levels of C. glutamicum cellular mRNA. Such

XX monitoring particularly differentiates between expression levels of

XX different strains of C. glutamicum and allows the adjustment of different

XX glutamicum derived polynucleotides described in the disclosure of the

XX invention

XX Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;

XX Query Match 14.1%; Score 106; DB 7; Length 177;

XX Best Local Similarity 100.0%; Pred. No. 1.5e-22;

XX Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGCCCTG 60

DB 72 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGCCCTG 131

QY 61 GAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTTC 106

DB 132 GAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTTC 177

RESULT 13

AAAF61688/C

AAAF61688 standard; DNA; 715 BP.

XX AAFA61688;

XX 12-JUN-2001 (first entry)

XX C. glutamicum lrp encoding DNA.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;

XX medicine; animal feed supplement; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers

XX -35_signal 62..67

XX -10_signal 88..93

XX CDS 151..615

XX /tag= a

XX /tag= b

XX /tag= c

XX /product= "lrp"

XX EP1090993-A1.

XX 11-APR-2001.

XX 29-SEP-2000; 2000EP-00121159.

XX	PR	05-OCT-1999;	99DE-0104792.
XX	PA	(DEGS) DEGUSSA-HUELS AG.	
PI	PI	Moeckel B, Péffertle W, Puenher A, Kalinowski J, Bathe B;	
DR	DR	WPI; 2001-292927/31.	
XX	DR	P-PSDB; AAB70881.	
PT	PT	New lrp gene from coryneform bacteria, used to prepare transformants with	
XX	PT	increased synthesis of amino acids, particularly lysine and isoleucine.	
PS	PS	Claim 4; Page 14-15; 22pp; German.	
XX	CC	This invention describes a novel isolated nucleic acid (I) from	
XX	CC	coryneform bacteria used for the fermentative production of selected L-	
XX	CC	amino acids, by fermenting the amino acid-producing coryneform in which	
XX	CC	at least the lrp gene has been weakened or amplified, then isolating	
XX	CC	c amino acids that have accumulated in the medium or cells. (I) is used to	
XX	CC	transform coryneforms for production of L-amino acids, specifically	
XX	CC	lysine and isoleucine, which are used in medicine and particularly as	
XX	CC	animal feed supplement. It may also be used as probes and primers for	
XX	CC	isolating related sequences. Regulating expression of (I) improves	
XX	CC	production of amino acids, especially of L-lysine. This sequence encodes	
XX	CC	the Corynebacterium glutamicum lrp protein which is used in the method	
XX	CC	described in the invention	
SQ	SQ	Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;	
		Query Match	11.3%; Score 85; DB 4; Length 715;
		Best Local Similarity	100.0%; Pred. No. 9.6e-16;
		Matches	85; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	OY	1 GTGCAAAAACGGAGAAGATTCAATTCAGCCTCGAGGTGTGCCATCCAAGCAGCCCTG	60
DB	DB	85 GTGCAAAAAACGAAAGATTCATTCGAACCTCGAGGTGTGCCATCCAAGCAGCCCTG	26
OY	OY	61 GAACCAGATGATAAAGTTATTCGGC	85
DB	DB	25 GAACCAGATGATAAAGTTATTCGGC	1
<hr/>			
RESULT 14			
ADA71938/c			
ID	ID	ADA71938 standard; DNA; 2000 BP.	
XX	AC	ADA71938;	
XX	DT	20-NOV-2003 (first entry)	
XX	DE	Rice gene, SEQ ID 5263.	
XX	KM	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	KM	gene; ds.	
XX	OS	Oryza sativa.	
XX	PN	MO2003000898-A1.	
XX	PD	03-JAN-2003.	
XX	PF	22-JUN-2001; 2001WO-IB001105.	
XX	PR	22-JUN-2001; 2001WO-IB001105.	
XX	PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX	PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
XX	PI	Kategirai F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;	
XX	DR	WPI; 2003-175290/17.	

Query Match	6.2%	Score 47	DB 7	Length 2000
Best Local Similarity	10.2%	Pred. No. 0.001		
Matches	75	Conservative 128	Mismatches 323	Indels 11, Gaps 2
XX	Claim 27	SEQ ID NO 5263	899bp	English.
XX				
CC	The present invention relates to a method (M1) for identifying genes			
CC	involved in plant resistance or response to pathogenic infection. M1			
CC	comprises identifying a gene whose expression is significantly altered in			
CC	the incompatible interaction of plant gene expression relative to			
CC	expression of the gene in an uninfected plant, in a mutant plant that			
CC	does not express a gene associated with response to pathogenic infection,			
CC	or in a corresponding incompatible or compatible interaction. (M1) is			
CC	useful for conferring resistance to resistance or tolerance to a plant to			
CC	bacterial, fungal or viral infection. The present sequence was used to			
CC	illustrate the invention.			
XX				
XX	Sequence 2000 BP, 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;			
QY	8	AAACGACAGAGATTCAATTCAGACCTCGAGAGTGTGCGCATCAAGACGACCTTGAAACAG	67	
DB	736	RMMSAMRYCSHMKCAKTKYASASAWTGRARSRFRMYRKRKGYTRYRWRSCMT	677	
QY	68	ATGATAAGGTTATGGGGGCTACGAAATCGCGCAAGGCTTAAACCTCCCTTGCGGAG	127	
DB	676	RAMSGRRKRAAGSMKSMCMWTRGRASMTYSKYSCKAOCKTRIMSTSMAGTGMTS	617	
QY	128	GTTTGGGACATGACCCGATGGTATGGTGGTCTCTTGGTTATTAATACGGCTACG	187	
DB	616	S--YKSMWTSKMSYMKATCTMTYSMKGSTRSRSMRMSGMSRMVYRMKKRKKRYAR	559	
QY	188	AATGGTGGGACGCCCACTGTTTTCCGGCGCATTTTGGGGGCGTCCACGAAATGCTGG	247	
DB	558	YMKRCKTRMCNCIRMGWMTTSSRMTYGRKRAATYSKRKYMYTKRYKTYTYGMA	499	
QY	248	TCATCGCCCT-----CGTTGGGCGCAGGCCCTGGGGCGCCATCGCGCTACCA	298	
DB	498	YMKCSYMMRGYCKACKCKCYAMCKAKAASGMMMYWTRKYSKMMRBSTKYSMMWYKOR	439	
QY	299	CATTGCTGTGAACCTCGGCCACGATTCATGAGGTTTCAATCCGCTGCATGTGATCA	358	
DB	438	SMKRGAGCYGCKMWTGYSYGMKYTTMSYKYSRCKTYRMTYGMWMTMYTAYSS	379	
QY	359	AAAACCCCATTTGCCGGTTCATTGCGTTTTGCGGCTTATCGAAGACCTACGACATCA	418	
DB	378	MMTWYYVYAKWYKWTYKRGATSMWYGKYSKRYCTWMCYMKCMCYMRDMRKKTRYK	319	
QY	419	CTGGGCGCAGGCCCGAGGCTGGTGGGGGTGGCGCACTATCTCAATCAATAGCGTTTC	478	
DB	318	RCTCMRYATYCCCTCRKRGWTSRSMSMRRTGAKMKMSMSMKCSYSHYMTYKMKKEY	259	
QY	479	ACTCTACTGGGTATTGGCGGCTCTACCCGAGTGGCGATCGCAGAATTGATTCCTTTTG	538	
DB	258	YMSYGNARSGTSTRSAAKRTYKYSTSRBRAMBRACGMSACRRVSRSTYCGCGSTCGS	199	
QY	539	AAATTTAAGGCGCCGAGTTGCGCCCTTTGCTCTCTCTTTGTACAGCTCACTTGAGATTC	598	
DB	198	SKTKMKSSKSMRMTSSSWCSGCTCYCYGAMCMSCSMWMTMSGCGTYTGMRKRSYSM	139	
QY	599	GCCGAAAGAAAGACGATCCTTCTCTGCTGCTCGCAGGTTGAGCTTCAACATTGCTC	658	
DB	138	CCKRYCSCKTKYCSYGYRYCKMYKYSYKYCYCYCMWYSYMRVYMKCMCSGCSGSM	79	
QY	659	TTGNGTAATTCAGGTCAGCCCATTTGCGGCGCTCATATCTTCTGGAGTGTGTA	718	
DB	78	SCAYCSTSSRSMRYTAAKMGCMGSGGMYRMSKCMVSKYSKVTYSKCTKRYK	19	
QY	719	CAATCGGACTTCTTTC	735	

Db 18 YCYWSSGYSNWCTSGTY 2

RESULT 15

ID AAA10594 standard; DNA; 10732 BP.

AC AAA10594 ;

DT 29-JUN-2000 (first entry)

DE Gene encoding a subunit of cellulose synthase

cellulose synthase; cellulose production; increase yield; ds

OS Vigna angularis.
YY

PN JP2000060568-A.
YY

PD 29-FEB-2000
XX

XX 46-AUG-1998; 9802-00239998.

26-AUG-1998; 9805-00239998.
FK
XX

PA (MIZU/) MIZUNO K.
PA (OJIP) OJI PAPER CO.

AA
DR
WPI: 2000-342371/30.

F-5000, 44402412.
 DN
 XX

the amount of cell

PS Claim 2; Page 14-21; 32pp; Japanese.

CC This sequence represents a gene encoding a subunit c

cellulose synthetic equipment, that

CC can also be used to improve the properties of the cellulose being

XX

Best Local Similarity 14.7%; Pred. No. 0.13;

.....

[illegible][illegible][illegible][illegible]

9829 SDSTSTVRCSPDYDATBSQDNSTNCYDASRTBTBSTNCYARCYMBYDARCSRDNSTVSSR 977

442 TCGGCGTGGCGACTTATCTCAATGCAATAGCGTTTCACTCCTACTGGGTAATCGGCGGT 501

db 9769 GYDANSTSRYSSTYSSDSTYSACKCAKSTBTBCYYPAYDACYDANCYSSDST 9771

502 CTCACCGAGTGGCGATCGCAGAGTTGATTCTTTTGAATTAAAGGCCCTCG-AGTTCCG 560

Db 9709 YTBYSRRCYYDAYSCSRYDARCYDACYSNSTCYDATBTSTYSYSTYSNCYDA 965

561 CCTTGGCTCTCTCTTTGTGACGCTGACTTTGGATTCTCGCCGAACGAAAGAAGCAGATCCC 620

Db 9649 TTSRCTBYSTBTTSRCACCTBDSTAKNSTSTYSTRCTBYSSRSRGYSYCSRSSR 955

```

Qy      621  TTCCTGCTGCTGCACAGTTTGACTTTCACCATGCTCTTGTGGTAAATCCAGGCTCAGGC   680
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      959  CACGCTDSDSTCVCSTTSTTAAVASCCTSRGVSXDASSTSTYSTSCTSTSTVSYSTTDYDC   9530

Qy      681  CCTATTGCGGCGGCGTG   696
      : : : : : : : : : :
Db      9529  SDYSTTBNSTYSDS   9514

```

Research completed: April 19 2004 11:42:00

Job time : 352.669 secs

Tue Apr 20 06:47:16 2004

us-10-608-504-2.rnt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 63.4023 Seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-2

Perfect score: 1 gtcgcaaaacgcaagagat.....ctctgggaaagctcgtctaa 753

Sequence: 1 gtcgcaaaacgcaagagat.....ctctgggaaagctcgtctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/pdata/2/ina/5A.COMB.seq:*
2: /cgn2_6/pdata/2/ina/5B.COMB.seq:*
3: /cgn2_6/pdata/2/ina/6A.COMB.seq:*
4: /cgn2_6/pdata/2/ina/6B.COMB.seq:*
5: /cgn2_6/pdata/2/ina/CTUS.COMB.seq:*
6: /cgn2_6/pdata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	US-09-471-803A-2	Sequence 2, Appl1
2	753	100.0	1271	US-09-471-803A-1	Sequence 1, Appl1
3	753	100.0	1271	US-09-471-803A-6	Sequence 6, Appl1
4	58.6	7.8	756	US-09-489-039A-3016	Sequence 3016, Ap
5	46.2	6.1	7218	US-08-232-463-14	Sequence 14, Appl
6	39.8	5.3	768	US-09-107-532A-1266	Sequence 1266, Ap
7	39.2	5.2	1038	US-09-252-991A-11376	Sequence 11376, A
8	39.2	5.2	1086	US-09-252-991A-11376	Sequence 11418, A
9	39.2	5.2	1086	US-09-252-991A-11376	Sequence 8976, Ap
10	36.8	4.9	1140	US-09-252-991A-3205	Sequence 3205, Ap
11	36.8	4.9	2061	US-09-252-991A-2761	Sequence 2761, Ap
12	36.4	4.8	751	US-09-020-956-12	Sequence 12, Appl
13	36.4	4.8	751	US-09-030-607-12	Sequence 12, Appl
14	36.4	4.8	751	US-09-439-313-12	Sequence 12, Appl
15	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appl
16	36.4	4.8	751	US-09-232-149A-12	Sequence 12, Appl
17	36.4	4.8	751	US-09-159-812-12	Sequence 12, Appl
18	36.4	4.8	751	US-09-636-215-12	Sequence 12, Appl
19	36.4	4.8	751	US-09-685-166A-12	Sequence 12, Appl
20	36.4	4.8	751	US-09-115-453-12	Sequence 12, Appl
21	36.4	4.8	751	US-09-688-489-12	Sequence 12, Appl
22	35.4	4.7	430	US-09-621-976-16656	Sequence 16656, A
23	35.4	4.7	474	US-09-621-976-16656	Sequence 18033, A
24	35.2	4.7	801	US-09-328-352-2389	Sequence 2389, A
25	35.2	4.7	4403765	US-09-103-840A-2	Sequence 2, Appl1
26	35.2	4.7	4411529	US-09-103-840A-2	Sequence 1, Appl1
27	34.8	4.6	867	US-09-489-039A-5047	Sequence 5047, Ap

28	34.6	4.6	666	2	US-08-875-034A-1	Sequence 1, Appl1
29	34.4	4.6	744	4	US-09-489-039A-1307	Sequence 1307, Ap
30	34.4	4.6	1224	4	US-09-266-965-22	Sequence 22, Appl
31	34.4	4.6	12249	4	US-09-266-965-74	Sequence 74, Appl
32	34.4	4.6	18311	4	US-09-266-965-96	Sequence 96, Appl
33	33.4	4.4	1875	3	US-09-422-869-21	Sequence 21, Appl
34	33.4	4.4	2949	3	US-08-433-522A-1	Sequence 1, Appl1
35	33.4	4.4	2949	3	US-09-135-166-1	Sequence 1, Appl1
36	33.4	4.4	2949	3	US-08-942-046-1	Sequence 1, Appl1
37	33.4	4.4	2950	3	US-08-433-522A-5	Sequence 5, Appl1
38	33.4	4.4	2950	3	US-09-135-166-5	Sequence 5, Appl1
39	33.4	4.4	2950	3	US-08-942-046-5	Sequence 5, Appl1
40	33.4	4.4	2984	3	US-08-433-522A-3	Sequence 3, Appl1
41	33.4	4.4	2984	3	US-09-135-166-3	Sequence 3, Appl1
42	33.4	4.4	2984	3	US-08-942-046-3	Sequence 3, Appl1
43	33.4	4.4	2987	3	US-08-433-522A-55	Sequence 55, Appl
44	33.4	4.4	2987	3	US-09-135-166-55	Sequence 55, Appl
45	33.4	4.4	2987	3	US-08-942-046-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1	US-09-471-803A-2
Sequence 2, Application US/09471803A	
Patent No. 6613545	
GENERAL INFORMATION:	
APPLICANT: KENNERKNECHT, NICOLE	
APPLICANT: SAMM, HERMANN	
APPLICANT: EGGELING, LOTMAR	
APPLICANT: PFEFFERLE, WALTER	
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF	
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE	
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF	
FILE REFERENCE: 21123/265496/MAS	
CURRENT APPLICATION NUMBER: US/09/471,803A	
PRIOR FILING DATE: 1999-12-23	
PRIOR APPLICATION NUMBER: DE 199 51 708-8	
PRIOR FILING DATE: 1999-10-27	
NUMBER OF SEQ ID NOS: 12	
SOFTWARE: PatentIn Ver. 2.1	
SEQ ID NO 2	
LENGTH: 753	
TYPE: DNA	
ORGANISM: Corynebacterium glutamicum	
FEATURE:	
NAME/KEY: CDS	
LOCATION: (1) ..(753)	
OTHER INFORMATION: BrnF	
OTHER INFORMATION: ATCC14752	
US-09-471-803A-2	
Query Match	100.0%; Score 753; DB 4; Length 753;
Best Local Similarity	100.0%; Pred. No. 2.2e-231;
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATTCAGGACGACCTG 60
DB	1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATTCAGGACGACCTG 60
QY	61 GAACCAATGATTAAGTTACGCGCTACGAAATCGCGAAGGTCTAAACCTCCCTT 120
DB	61 GAACCAATGATTAAGTTACGCGCTACGAAATCGCGAAGGTCTAAACCTCCCTT 120
QY	121 GCTGAGGTTGGGACATGACCCGATTTGATTCGTTGGTCTCTTTGTTATTCATAC 180
DB	121 GCTGAGGTTGGGACATGACCCGATTTGATTCGTTGGTCTCTTTGTTATTCATAC 180
QY	181 GGCTACGATGATGAGGAGGAGCCGCTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
DB	181 GGCTACGATGATGAGGAGGAGCCGCTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240

QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 300
 Db 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 300
 QY 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 360
 Db 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 360
 QY 361 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 420
 Db 361 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCACACA 420
 QY 421 GCGGCGAGCCCGCAGGCTGTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 480
 Db 421 GCGGCGAGCCCGCAGGCTGTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 480
 QY 481 TCTCTAGGTAATTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 540
 Db 481 TCTCTAGGTAATTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 540
 QY 541 ATTAAGGCTCGAGTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 600
 Db 541 ATTAAGGCTCGAGTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 600
 QY 601 CGAAGCAAAAGAGATCCCTTCTGCTGCTCGAGGTTGAGCTTCACATTCCTCT 660
 Db 601 CGAAGCAAAAGAGATCCCTTCTGCTGCTCGAGGTTGAGCTTCACATTCCTCT 660
 QY 661 GTGGTAATTCAGGTCAGGCTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 720
 Db 661 GTGGTAATTCAGGTCAGGCTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 720
 QY 721 ATCCGTAATTCCTTCTGCGAAGGCTGCTAAA 753
 Db 721 ATCCGTAATTCCTTCTGCGAAGGCTGCTAAA 753

RESULT 2

US-09-471-803A-1

; Sequence 1, Application US/09471803A
 ; Patent No. 6613545
 ; GENERAL INFORMATION:
 ; APPLICANT: KENNERNKUCHT, NICOLE
 ; APPLICANT: SAHM, HERMANN
 ; APPLICANT: PFEFFERLE, WALTER
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 ; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 ; FILE REFERENCE: 21123/265496/MAS
 ; CURRENT APPLICATION NUMBER: US/09/471, 803A
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: DE 199 51 708.8
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1271
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (110)..(853)
 ; OTHER INFORMATION: DTMF
 ; NAME/KEY: gene
 ; LOCATION: (853)..(1176)
 ; OTHER INFORMATION: DTMF
 ; OTHER INFORMATION: ATCC14752
 ; US-09-471-803A-1

Query Match 100.0%; Score 753; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 2,9e-21;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAGAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGGACGCTG 60
 Db 101 GTGCAAAAAGCAGAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGGACGCTG 160
 QY 61 GAACAGATGATAAGGTTATCGGCGCTACGAATCGCGCAAGGTCATAAAACCTCCCTT 120
 Db 161 GAACAGATGATAAGGTTATCGGCGCTACGAATCGCGCAAGGTCATAAAACCTCCCTT 220
 QY 121 GCTGCAAGTTTGGGCAATGACCGAATGGTATGGTGTGCTGTGGTTATTCATAC 180
 Db 221 GCTGCAAGTTTGGGCAATGACCGAATGGTATGGTGTGCTGTGGTTATTCATAC 280
 QY 181 GGTACGAATGATGAGGACGCGCAGCTGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 240
 Db 281 GGTACGAATGATGAGGACGCGCAGCTGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 340
 QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 300
 Db 341 ATGCTGTCATCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 400
 QY 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 360
 Db 401 TTGCTGTCATCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 460
 QY 361 AACCCATTCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 420
 Db 461 AACCCATTCGCCCTGTTGTGGGCGCAGGCGCCCTGAGGCGCAATGCGGCTCACACA 520
 QY 421 GCGGCGAGCCCGCAGGCTGTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 480
 Db 521 GCGGCGAGCCCGCAGGCTGTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 580
 QY 481 TCTCTAGGTAATTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 540
 Db 581 TCTCTAGGTAATTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 640
 QY 541 ATTAAGGCTCGAGTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 600
 Db 641 ATTAAGGCTCGAGTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 700
 QY 601 CGAAGCAAAAGAGATCCCTTCTGCTGCTCGAGGTTGAGCTTCACATTCCTCT 660
 Db 701 CGAAGCAAAAGAGATCCCTTCTGCTGCTCGAGGTTGAGCTTCACATTCCTCT 760
 QY 661 GTGGTAATTCAGGTCAGGCTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 720
 Db 761 GTGGTAATTCAGGTCAGGCTTCGCGCGTGGCACTTAATCGAAGCCATCGAGTCACT 820
 QY 721 ATCCGTAATTCCTTCTGCGAAGGCTGCTAAA 753
 Db 821 ATCCGTAATTCCTTCTGCGAAGGCTGCTAAA 853

RESULT 3

US-09-471-803A-6

; Sequence 6, Application US/09471803A
 ; Patent No. 6613545
 ; GENERAL INFORMATION:
 ; APPLICANT: KENNERNKUCHT, NICOLE
 ; APPLICANT: SAHM, HERMANN
 ; APPLICANT: PFEFFERLE, WALTER
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 ; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 ; FILE REFERENCE: 21123/265496/MAS
 ; CURRENT APPLICATION NUMBER: US/09/471, 803A
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: DE 199 51 708.8
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patentin Ver. 2.1

Tue Apr 20 06:47:16 2004

us-10-608-504-2.rnt

Page 3

SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnF
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnE
OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,9e-231;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCGAAGATTCATTCAAGCCTGAGAGGTGGCCATCCAGGAGGCGCCCTG 60
DB 101 GTGCAAAAAGCGAAGATTCATTCAAGCCTGAGAGGTGGCCATCCAGGAGGCGCCCTG 160
QY 61 GAACCAAGATGATTAAGTTATCGGCGTACGAAATCGCGAAGGTCTAAAACCTCCCTT 120
DB 161 GAACCAAGATGATTAAGTTATCGGCGTACGAAATCGCGAAGGTCTAAAACCTCCCTT 220
QY 121 GCTGCAAGTTGGGCGATGTAACCCGATTTGGTATGGTCTTCTTGGTTATTCATAC 180
DB 221 GCTGCAAGTTGGGCGATGTAACCCGATTTGGTATGGTCTTCTTGGTTATTCATAC 280
QY 181 GCTGCAAGTTGGGCGATGTAACCCGATTTGGTATGGTCTTCTTGGTTATTCATAC 240
DB 281 GCTGCAAGTTGGGCGATGTAACCCGATTTGGTATGGTCTTCTTGGTTATTCATAC 340
QY 241 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTCGGGCGCATGCGCTCACCA 300
DB 341 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTCGGGCGCATGCGCTCACCA 400
QY 301 TTGCTGAGTGAATTCGCGCGACGATTTCTATGCGTTTCAATCGCGCTGATGGTCAAA 360
DB 401 TTGCTGAGTGAATTCGCGCGACGATTTCTATGCGTTTCAATCGCGCTGATGGTCAAA 460
QY 361 AACCCGATTTGCGGCTTCTATTCGTTTTCGCGCTTATGACGAAAGCTTACGACTACT 420
DB 461 AACCCGATTTGCGGCTTCTATTCGTTTTCGCGCTTATGACGAAAGCTTACGACTACT 520
QY 421 GCGGCGAGGCGCGGAGGCTGCTGCGGCTGCGGCTTATTCGAAATGCGGTTTAC 480
DB 521 GCGGCGAGGCGCGGAGGCTGCTGCGGCTGCGGCTTATTCGAAATGCGGTTTAC 580
QY 481 TCTCTATGAGTATTCGCGGCTTCTACCGGAGTGGCGATGCGAGTTGATTCCTTTGA 540
DB 581 TCTCTATGAGTATTCGCGGCTTCTACCGGAGTGGCGATGCGAGTTGATTCCTTTGA 640
QY 541 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTTCTTGTACCGTGAAGTTGATTCCTGCG 600
DB 641 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTTCTTGTACCGTGAAGTTGATTCCTGCG 700
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCATGCTCTT 660
DB 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCATGCTCTT 760
QY 661 GTGTAATTCAGTGAAGGCGCTTATTTGGGCGCTGCTGATCTTCTTGGGCTCTGTTGAC 720
DB 761 GTGTAATTCAGTGAAGGCGCTTATTTGGGCGCTGCTGATCTTCTTGGGCTCTGTTGAC 820
QY 721 ATCCGAGTCTTCTCTTGGGAAAGGCTGCTAAA 753
DB 821 ATCCGAGTCTTCTTGGGAAAGGCTGCTAAA 853

RESULT 4
US-09-489-039A-3016

Sequence 3016, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 3016
LENGTH: 756
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

Query Match 7.8%; Score 58.6; DB 4; Length 756;
Best Local Similarity 52.2%; Pred. No. 8.6e-09;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 95 TCGCGAAGGCTTAAGAACTCCCTTCTGCTGCAAGTTGGGCGATGTAACCCGATTTGATG 154
DB 62 TCGCGAAGGCTTAAGAACTCCCTTCTGCTGCAAGTTGGGCGATGTAACCCGATTTGATG 121
QY 155 GCTTGGTCTCTGTTATTCATACGCGTACGATGTTGGGCGAGCCGCACTGTTTTCG 214
DB 122 CTTGCGACATGATGTAACCCGCTGCTTACCCCGCGTAAGACGTTCTTCTCT 181
QY 215 GCTGATTTTCGCGGCTTCCAGCAATGCTGCTATCGCCCTGTTGGGCGAGCCG 274
DB 182 GCATTAATTCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241
QY 275 CCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 334
DB 242 CCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
QY 335 TTTCATTCG 343
DB 302 CTTCACTGC 310

RESULT 5
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: SCHRIFFLER, F.
APPLICANT: DOERNER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-9300
 TELEFAX: (703) 683-4109
 TELEX: 899149
 INFORMATION FOR SEQ. ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZpdt-Fls
 US-08-232-463-14

Query Match 6.1%; Score 46.2; DB 1; Length 7218;
 Best Local Similarity 5.2%; Pred. No. 0.00027;
 Matches 21; Conservative 214; Mismatches 172; Indels 0; Gaps 0;

215 GCGTATTTTCGGGGCTCCACGGAATGCTGTATGCGCCCTGTTGGGCGAGCGC 274
 1041 GGGTGAAGCTCGAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTTTTTT 1100
 275 CCCTGGGCGCCATCGCGTACCAACATGCTGTGTAATTCGCGCAATTTCTATGCGT 334
 1101 YY 1160
 335 TTTCATTCGCGCTGATGCTGCAAAAACCCATTCGCGCTTCTATGCTTTGCGCC 394
 1161 YY 1220
 395 TTATCGAGAGCTGATCGCATGCGGCGAGCGCCGCGAGGCTGCGCGTGGCGAGC 454
 1221 YY 1280
 455 TTATTCATGCAATAGCTTTTCACTCTCTACTGCTATTCGGCGCTTCACCGAGTGG 514
 1281 YY 1340
 515 CGATCGAGAGTGTATGCTTTGGAATTAAGGCGCTCGAGTTCGCGCTTCTCTCT 574
 1341 YY 1400
 575 TTGTACGCTGACTTTGATTCCTCGCGAGCAAAAACGATCCCT 621
 1401 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAATTCCT 1447

RESULT 6
 US-09-107-532A-1266
 Sequence 1266, Application US/09107532A
 Patent No. 6583275

GENERAL INFORMATION:
 APPLICANT: Lynn A. Doucetite-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: m1ec_feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-09-107-532A-1266

Query Match 5.3%; Score 39.8; DB 4; Length 768;
 Best Local Similarity 4.83%; Pred. No. 0.0092;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

107 TAAACCTCCCTGCTGCAAGTTTGGCATGTAACCGATTTGATTTGCTCT 166
 92 TCAAGACAGCTACTACCTGTTTCGTTATATGCTATTTGACATTGGATTCG 151
 167 TGTATTATCATATAGGCTATGATGAGGAGGCGGCACTGTTTCCGGCTGATTTCC 226
 152 TTGGAAAGCTGCGGATTTATCACTATGCTGACCTGATGCTCTGCTGATG 211
 227 CGGCTCCACGGAATGCTGTCATCGCCCTCGTTGTGGGCGAGCGCCCTGGGCGCCA 286
 212 CTGTTTGGCCCAATTATACAGTACAGTACGATGCTGCTATGACAGCCCAATGCTTCCA 271
 287 TCGGCTACCAACATGCTGTGTAATTCG 317
 272 TCGTTTCTCGACCTTCTAGTCAATTCGCG 302

RESULT 7
 US-09-252-991A-11376
 Sequence 11376, Application US/09252991A
 Patent No. 6551795

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 11376
 LENGTH: 1038
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Query Match 5.2%; Score 39.2; DB 4; Length 1038;
 Best Local Similarity 55.0%; Pred. No. 0.017;
 Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

209 TTTCGGCGGCTATTTGGCGGGCTCCACCGAATGCTGATCGCCCTGTTGCGCG 268
 Db TGTCCGCGATGCTTTCGCGCGCGCCGCGCCGCGGCGGCGGCGATGCTCAAGGCG 532
 473 TGTCCGCGATGCTTTCGCGCGCGCCGCGCGGCGGCGGCGGCGATGCTCAAGGCG 532
 QY CAGCGCGCGGCTGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 328
 Db GTGCGCGGCTCTTTTCGATCACTTCAACCCCTCTCTGACCTCGAGAGAGAGAGAG 592
 QY 329 ATGCGTTTCATTTCCGCGT 348
 Db 593 ACGGATGAGCTTGGCGCG 612

RESULT 8

US-09-252-991A-11418
 ; Sequence 11418, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 11418
 LENGTH: 1086
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-11418

Query Match 5.2%; Score 39.2; DB 4; Length 1086;
 Best Local Similarity 55.0%; Pred. No. 0.017;
 Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

209 TTTCGGCGGCTATTTGGCGGGCTCCACCGAATGCTGATCGCCCTGTTGCGCG 268
 Db 222 TGTCCGCGATGCTTTCGCGCGCGCCGCGCGGCGGCGGCGGCGGCGGCGGCGG 281
 QY CAGCGCGCGGCTGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 328
 Db 282 GTGCGCGGCTCTTTTCGATCACTTCAACCCCTCTCTGACCTCGAGAGAGAGAG 341
 QY 329 ATGCGTTTCATTTCCGCGT 348
 Db 342 ACGGATGAGCTTGGCGCG 361

RESULT 9

US-09-621-976-8976/c
 ; Sequence 8976, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Ubert, S.
 TITLE OF INVENTION: ESTE and Encoded Human Proteins.
 FILE REFERENCE: GENSET.054PR2
 CURRENT APPLICATION NUMBER: US/09/621,976
 CURRENT FILING DATE: 2000-07-21
 NUMBER OF SEQ ID NOS: 19335
 SOFTWARE: Patent.pm
 SEQ ID NO 8976

LENGTH: 399
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-621-976-8976

Query Match 5.2%; Score 39; DB 4; Length 399;
 Best Local Similarity 12.6%; Pred. No. 0.012;
 Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

QY 509 GAGTGGCGATCCGAGAGTTCCTTTGAAATTAAGGCGCTCGAGTTCGCTTGT 568
 Db 242 SAMRRKRRKMTCKMGRSWSRSTGYAMWKSWCTSKRMWYKRRKRRKSTRT 183
 QY CTCTCTTTGTCACGCTGATCTTGGATTCCTGCGGAGAAAAAGAGATCCTTCTGC 628
 Db 182 CYRSTYKCMKAYTKRRKRRWYTYKYSVSMKTRMTAYYWTGRWTRTKMTW 123
 QY 629 TGTCTCGAGTTGAGCTTCACCATTCCTGTGTATTCAGTCCAGGCGGCTATT 688
 Db 122 CTMCKCTTYMAGTMYRRYRRYAKAKMSKCTWSTTCYCMKYMCKMSTWMS 63
 QY 689 CGGCGCTGCTGATCTTCTGCTGTGTGACCATCCGACTCTTCTTGGAGAGGCT 747
 Db 62 MMCKGSMWKKWYTYTYTYMCKMSTWMSVYARKCMRYAKTYMTCTWCKRT 4

RESULT 10

US-09-252-991A-3205/c
 ; Sequence 3205, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 3205
 LENGTH: 1140
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3205

Query Match 4.9%; Score 36.8; DB 4; Length 1140;
 Best Local Similarity 58.0%; Pred. No. 0.1;
 Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 211 TCGCGCTGATTTTCGCGGGCTCCACCGAATGCTGATCGCCCTGTTGCGGCA 270
 Db 465 TACTTCTGATCACTTCGCGCGCGCGCGCGCGCGCGCGCGGCTGATCGGCTTCCGCGTC 406
 QY 271 GCGCGCTGCG 322
 Db 405 GCG 354

RESULT 11

US-09-252-991A-2761
 ; Sequence 2761, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 4.8%; Score 36.4; DB 4; Length 2061;
Best Local Similarity 58.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 47;

QY 211 TCCGCGCTGATTTTCGCGGCTCCACCAATGCTGTCATCGCCCTGTTGGGCGCA 270
DB 1273 TACTTCTGATCAACTCGGCGCGCCCTCGGCGCGCTGATGCGCTGACTCGGCGCTC 1332
QY 271 GCGCCCTCGGCGCGCTCCGCGCTCACACATGCTGCTGAATTCGCGCAGC 322
DB 1333 GCGCGCCAGCGCGCGCTTCTCGTCAACGCGCGCTGCTACTTCTGCTACG 1384

RESULT 12
US-09-020-956-12/C
Sequence 12, Application US/09020956
Patent No. 6261562

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-020-956-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; Indels 39; Gaps 0;
Matches 58; Conservative 0; Mismatches 39;

QY 227 CGGCGCTCAGCAATGCTGTCATCGCCCTGTTGGGCGCAGCGCCCTTGGGCGCA 286
DB 410 CGTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
QY 287 TCGGCGCTCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323

DB 350 TGGTGTACACCAATGCTGAGCACTTCCCTGACCT 314

RESULT 13
US-09-030-607-12/C
Sequence 12, Application US/09030607
Patent No. 6262245

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-030-607-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; Indels 39; Gaps 0;
Matches 58; Conservative 0; Mismatches 39;

QY 227 CGGCGCTCAGCAATGCTGTCATCGCCCTGTTGGGCGCAGCGCCCTTGGGCGCA 286
DB 410 CGTCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
QY 287 TCGGCGCTCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
DB 350 TGGTGTACACCAATGCTGAGCACTTCCCTGACCT 314

RESULT 14
US-09-439-313-12/C
Sequence 12, Application US/09439313
Patent No. 6329505

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yugu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John

```

? APPLICANT: Day, Craig
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
? TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
? FILE REFERENCE: 210121.427C9
? CURRENT APPLICATION NUMBER: US/09/439,313
? NUMBER OF SEQ ID NOS: 575
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 12
? LENGTH: 751
? TYPE: DNA
? ORGANISM: Homo sapien
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(751)
? OTHER INFORMATION: n = A,T,C or G
US-09-439-313-12

```


NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: mlec_feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-10-417-884A-1266

Query Match 5.3%; Score 39.8; DB 6; Length 768;
 Best Local Similarity 49.3%; Pred. No. 0.0082;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
 QY 107 TAAAAACCTCCCTGCTGAGGTTGGAGCATGATACCGATTGGTATGCGTTGGTCT 166
 DB 92 TCAAGACACGTTACTACCTGCTTTCGTTATGATGATGGACTTGGCAATTTGGTATCG 151
 QY 167 TGGTATTCATAGGCGCTAGAGATGGTGGGAGCCCACTGTTTCCGGCCGTATTTCC 226
 DB 152 TTGGAAAGCTGCCGATTTCAATCACTAGTCGTCACGTTATCCCTGGTGTCTATG 211
 QY 227 CGGGCTCCACGGAATGCTGTCATCGCCCTGTTGGGCGCAGCGCCCTGGGCGCCA 286
 DB 212 CTGGTTGCGCAATTTATCAAGTCAGTCAGCATGCTGTGATGTCAGAGCCCATTTGTTCCA 271
 QY 287 TCGCGCTCAGCAGATTGCTGTGGAACCTCCG 317
 DB 272 TCGTTTCTGACCTTTCAAGCAATTCCTCG 302

RESULT 3
 US-10-767-701-5263
 Sequence 5263, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5353)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 5263
 LENGTH: 619
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS11974_1
 US-10-767-701-5263

Query Match 4.7%; Score 35.4; DB 6; Length 619;
 Best Local Similarity 54.1%; Pred. No. 0.19; 61; Indels 0; Gaps 0;
 Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 226 GCGGCTCCACCGAATGCTGTCATGCGCTTCGTTGGGCGCAGCGCCCTGGGCGCC 285
 DB 132 GCGGCTTCACCTTGCTGCTGTGTCGCGGCTGGCGGTGACGGGCAAGTCTGGGCTC 191
 QY 286 ATCGGCTCCACCATGCTGCTGTAATTCGCGCAGTATTCATGCGTTTATTCCTCCG 345
 DB 192 ATCGTGGACGCGCTGCTGCTGATCTTCAAGCCCTGCTGCTGCTGCGGCGCATCAG 251
 QY 346 CTGATGCTGCTCA 358
 DB 252 GTGGCGCTGCTCA 264

RESULT 4
 US-10-767-701-10866
 Sequence 10866, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5353)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 10866
 LENGTH: 1034
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
 US-10-767-701-10866

Query Match 4.6%; Score 34.6; DB 6; Length 1034;
 Best Local Similarity 49.7%; Pred. No. 0.45; 89; Indels 0; Gaps 0;
 Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 QY 241 ATGCTGTATGAGCCCTGTTGTGGGCGAGCGCCCTGGGCGCATGCGCTCACACA 300
 DB 151 ATGGCGGCGCCCTACTACGCGGAGGCGCCCTGGGCGCACCGGCGCGCGCCCT 210
 QY 301 TTGCTGTGAATTCGCGCAGTATTCATGCGTTTATTTCCGCTGATGATGGTCAA 360
 DB 211 GTGGCGGTGTGTCCCGCAGTTTCGCGGCGGCTGAGTGTGCGCTTACGCTGACAG 270
 QY 361 AACCCATGCGCGTTTATTCGCTTATTCGCGCTTATGACGAGACCTACGCACTC 417
 DB 271 AAGGCATGAGCTTCTCCAGCGGCACTTCAACGTCACGAGCCAGGCGCGCTC 327

RESULT 5
 US-10-100-683-983/c

```
; Sequence 983, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-100-683-983

Query Match      4.6%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.79;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      244 CCGGTCATGCGCCCTGTTGGGCGGAGCGCCCTGGGGCGCATGCGGTCAACCA 300
DB      2360 CAGCTCCGCGCCCGCAGTGTGGGCTCAAGGCTCTGGGCGCATGAGTTCACCA 2304

RESULT 6
US-10-100-683-11511
; Sequence 11511, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
```

```
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 11511
; LENGTH: 7071
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-100-683-11511

Query Match      4.6%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.2;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY      244 CCGGTCATGCGCCCTGTTGGGCGGAGCGCCCTGGGGCGCATGCGGTCAACCA 300
DB      785 CAGCTCCGCGCCCGCAGTGTGGGCTCAAGGCTCTGGGCGCATGAGTTCACCA 841

RESULT 7
US-10-767-701-7545
; Sequence 7545, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 7545
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
; US-10-767-701-7545

Query Match      4.6%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.42;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY      215 GCTGATTTTGGGGGCTCCACCGAAATGCTGTATCGCTTCGTGTGGGCGGAGCCG 274
DB      31 GCCCGACACACACGAGAGAGAGACGCCGCTCATCAACCCGTTCTGGAGCGGCGCGC 90

QY      275 CCCGCGGCGCATGCGGCTGACCAATTGCTGGTGAATTCCGCCACG 322
DB      91 CGTGGCGTCTGCGGTGACGCGCGCTCTGTGTGCGGTGCGGAGG 138

RESULT 8
US-10-767-701-744/C
; Sequence 744, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 744
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
; US-10-767-701-744
```


Tue Apr 20 06:47:17 2004

us-10-608-504-2.rnpn

Page 6

Search completed: April 17, 2004, 20:12:58
Job time : 88.0151 secs

Tue Apr 20 06:47:17 2004

US-10-608-504-2.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 2438.38 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-2
Perfect score: 753
Sequence: 1 gtcgcaaaaacgcaagagat.....tcttggaagagctgctaaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pin:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rnd:*
26: em_ges_vrt:*
27: em_ges_vrt1:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.6	7.3	473	28	AQ936633 HSJ41-905
2	45	6.0	1201	13	BX356664 BX356664
3	39.4	743	29	CC684922	OGUAX55TH
4	39.4	5	787	29	CNS009WA
					AL054013 Drosophila

5	39.2	5.2	1201	13	BX381961	BX381961
6	39	5.2	925	29	CNS0091P	AL05013 Drosophila
7	38.6	5.1	1201	13	BX361080	BX361080
8	38	5.0	835	14	CB657749	OSUNEC13F
9	37.4	5.0	697	28	BH887860	BH887860 LB01712A.
10	37.4	5.0	1108	13	BX382084	BX382084 BX382084
11	37.2	4.9	1030	12	BG845919	BG845919 1024011FO
12	37.2	4.9	1454	28	BF101158	BF101158 60175473
13	36.8	4.9	1454	28	BH805539	BH805539 1008061AI
14	36.8	4.9	1201	29	CNS016BR	AL106545 Drosophila
15	36.6	4.9	539	10	BE013856	BE013856 125354 MA
16	36.6	4.9	942	29	CNS00CT2	AL059804 Drosophila
17	36.4	4.8	298	12	BJ525053	BJ525053 BJS25053
18	36.2	4.8	673	12	BJ287010	BJ287010 BJ287010
19	36.2	4.8	842	13	BX311708	BX311708 BX311708
20	36.2	4.8	902	29	CG358991	CG358991 OG3AU35TH
21	36.2	4.8	1101	13	BO678281	BO678281 AGRICOURT
22	36	4.8	559	13	BX384405	BX384405 BX384405
23	36	4.8	666	28	CC170529	CC170529 1197407.b
24	36	4.8	784	29	CG250542	CG250542 OGXE072TV
25	36	4.8	836	29	CG449984	CG449984 OGVE072TV
26	36	4.8	862	28	BZ644270	BZ644270 OGCRX45TC
27	36	4.8	863	29	CG449978	CG449978 OGVE072TV
28	35.8	4.8	410	9	AI138160	AI138160 TBEST23 T
29	35.8	4.8	641	28	BZ375719	BZ375719 1e60910.9
30	35.8	4.8	641	28	CC668174	CC668174 OGULY96TV
31	35.8	4.8	696	29	CC654373	CC654373 OGVEH12TV
32	35.8	4.8	980	29	CG326716	CG326716 OG2C063TV
33	35.6	4.7	327	10	BF543765	BF543765 UI-R-C3-t
34	35.6	4.7	325	13	BX424877	BX424877 BX424977
35	35.6	4.7	532	9	AI293892	AI293892 LB07146.5
36	35.6	4.7	657	13	CA109167	CA109167 SCGSR106
37	35.6	4.7	683	14	CF850119	CF850119 pEMu013XA
38	35.6	4.7	694	28	CC337836	CC337836 OGTAFL14TV
39	35.6	4.7	798	12	BG855128	BG855128 1024041HO
40	35.6	4.7	900	29	CG371461	CG371461 OG2A005TV
41	35.6	4.7	977	29	CNS00UX7	AL076880 Drosophila
42	35.6	4.7	985	13	BX359916	BX359916 BX359916
43	35.6	4.7	1201	13	BX381971	BX381971 BX381971
44	35.4	4.7	398	14	CF865113	CF865113 pZS014XB
45	35.4	4.7	454	9	AI575130	AI575130 UI-R-G0-u

ALIGNMENTS

RESULT 1
AQ936633
LOCUS HSJ41-905 Human Noci clones Homo sapiens genomic, genomic survey
DEFINITION
ACCESSION AQ936633
VERSION AQ936633.1 GI:7213011
KEYWORDS
SOURCE GSS.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Zabarovsky, E.R., Gazatullin, R., Podowski, R.M., Zabarovsky, V.V., Xie, L., Miravento, O.V., Kozlyev, S., Petrenko, L., Skobeleva, N., Li, J., Protodopov, A., Kashuba, V., Ernberg, I., Winberg, G. and Wahlstedt, C.
Noci clones in the analysis of the human genome

TITLE
JOURNAL Nucleic Acids Res. 28 (7), 1635-1639 (2000)
MEDLINE
PUBMED
COMMENT
Contact: Podowski RM
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337983

Tue Apr 20 06:47:17 2004

us-10-608-504-2.rst

Page 2

FEATURES

Email: Raf.Podowski@cgr.ki.se
Class: NotI site.
Location/Qualifiers

ORIGIN

Query Match	7.3%	Score 54.6	DB 28	Length 473
Best Local Similarity	56.2%	Pred. No. 0.00068		
Matches 99; Conservative	0;	Mismatches 77;	Indels 0;	Gaps 0;

QY 123 TGGGAGAGCCCACTGTTTTCGAGCCCGATTTTCGCGAGCTCCAGCGAAATGCTGATC 252

Db 223 TGGAGAGGCGACAGGAGCTCTGGGCATCGTCTTTCGCGGGGCGCGCGAAGTTGGTGCATC 342

QY 223 GCCCTCTGTTGTGGGCGACAGCGCCCTGGGGCGCATCGGCTTACACATGCTGGGAGAC 312

Db 343 GGCATGCTTAAAGGCGAGGCGACCACTGCTGTGATCTCTGACACACATGCTGCTGACT 402

QY 313 TTCCGACACGATTTCTATGCGTTTTCATTCGCCGTCATGTGGTCAAAAACCCCAT 368

Db 403 TCGACGACACTCTTATTACGAGTGTTCATGCGCGCTGTGCTGACAGGCAACCCCTT 458

RESULT 2	
BX356664	1201 bp mRNA linear EST 05-MAY-2003
LOCUS	BX356666 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION	clone CSOD015YB03.3-PRIME, mRNA sequence.
ACCESSION	BX356664
VERSION	BX356664
KEYWORDS	BX356664.1 GI:30378083
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

ORIGIN

Query Match	6.0%;	Score 45;	DB 13;	Length 1201;
Best Local Similarity	12.9%;	Pred. No. 0.48;		
Matches 47;	Conservative 168;	Mismatches 148;	Indels 0;	Gaps 0

```

Oy      207 GATTTCGGGCTCATTTTCCGGGGCTCCACCGAATGCGGTACGCCCTCGTTGAGG   266
          :|::||:::||||:::||||:::||||:::||||:::||||:::||||:::
Db      745 TATSTASTTBTBSCITTBSTBTBTBBSYBABSBBTTSBCSTSSBSTTSTBTBMMS   804

```

QY	267	CGCAGCGCCCTTGAGGGCCCATCGCGCTCACACACATGTCGTGAACCTCGGCACCGATT	322
Db	805	SBCTCTSSSSSSSSSTTSTSTTSTTSKSTSTTSBBTBSSSSSTSTWTBTCTCTSTWTTS	864
QY	327	CTATGCGCTTTTCATCTCCCGCTGCATGTGTCAAAAACCCCATTTGCCCGTTTCTATTCGT	386
Db	865	YSSBSTBTBSTSTSTBBSwSBHTTTTBTBTBTWTSSSTSTTBSSSSSSSSSTTWS	924
QY	367	TTTGCGCTTATGCACGACGCGCTACGAGTCACTCGCGACGCGCCCGACAGCTGTGCGC	446
Db	925	BTSSSSSTTSSSSSSTBTSSSSSSSSSBTSSKSSBSBSSSSSSSTBTSTTTT	984
QY	447	GTGCGCACTTATCTCAATGCAAATAGCGTTTCACTCTACTGGGTATTGCGCGGTCTCAC	506
Db	985	VTSSBRTTSTBTBTBBSSTTTSSSTBSSBBSSTSSSSSSSSSTBTSSSNKSSSSS	1044
QY	507	CGAGTGCACATCGCAGAGTGATCTCTTTTGAATTAAGGCGCTCGAGATGCGCCCTTTG	566
Db	1045	SSSSSSSSSSSSSSSBTTTBTBTBTBBSSSSBSTTTTSSSSSSSSTBTSSSTTTT	1104
QY	567	CTC 569	
Db	1105	TTS 1107	

RESULT 3	CC684922	LOCUS	CC684922	743 bp	DNA	linear	GSS 19-JUN-2003
DEFINITION	CC684922	OGXAX55TH ZM.0.7.1.5_KB	Zea mays genomic clone ZMNM8a0397J14,				
ACCSSION	CC684922	Genomic survey sequence.					
VERSION	CC684922.1	GI:32089698					
KEYWORDS	GSS.						
SOURCE	Zea mays						
ORGANISM	Zea mays						

ORIGIN

Query Match	5.2%	Score 39.4;	DB 29;	Length 743;
Best Local Similarity	53.6%;	Pred. No. 13;		
Matches 82; Conservative	0;	Mismatches 71;	Indels 0;	Gaps 0

Oy 431 CCGAGGCTGTCGCGCGCTGGCGACTTATCTCAATGCAATAGCGTTTCACCTCTACTCGG 490
 Db 475 CTTCCGCTCCGCGCGCTTCGCTCCTCCACCTCTCACTCGGATTCGGCAATTTCTAGG 534

QY		491	TATTCGGGCGTCTACCCGAGGCGGATCGCAGAATTATTCCTTTAAATAAAGGCC	550
Dd		535	GTTTCGGCGCGGCACCACCGATCAACTGTGGCTTCCGACGACCTTAAGCTTGAGTC	594
QY		551	TCGAGTTGCGCCCTTGTCTCTCTCTTTGTACGC	583
Dd		595	ACGAGAGGCGCTTGTCTCTCTCTCTGTCGCCG	627
 RESULT 4 CONS009WA/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutomo Osoegawa and Aaron Mammeter in Pictet de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RpCl-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers 1..787 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACPAC20K04" /clone_1lb="RpCl-98" /note="end : T7"				
 ORIGIN				
Query Match		5.2%;	Score 39.4;	DB 29; Length 787;
Best Local Similarity		45.5%;	Pred. No. 14;	Mismatches 76; Indels 0; Gaps 0
Matches		76;	Conservative 15;	
QY		552	CGAGTTGCGCCTTGTCTCTCTTTGTACAAGCTATTGATTCCTGCGAACGAAGAAA	611
Dd		334	CTACTTCGTCCTTCTCTCTTTMOTGTATAMKMGKMKMTTGAATCCCKTKVAGARARAAC	275
QY		612	GGAGATCCCTTCTCTGCTGCTGCGCAGATTGAGCTCACCATGCTCTGTGGTATATCC	671
Dd		274	ATTGTGTTTTTTCACWGRGRCCTAAGGGGAATTCRTTAATTTTWTGTTTAATTT	215
QY		672	AGATCAGAGCCCTATTTCGCGCGCTGCTGATCTCTTTGGTCTGTGA	718
Dd		214	TAAATTCGCCATATAGTATTATATCGAATTTTTTTTATATGATCGCTGA	168
 RESULT 5 BX381961				

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
BX381961	1201 bp	RNA	linear	EST 08-MAY-2003							
BX381961	Homo sapiens PLACENTA COT 25-NORMALIZED	Homo sapiens cDNA									
clone CSOD1072YF05	3-PRIME, mRNA sequence.										
BX381961											
BX381961.1	GI:30453007										
EST.											
Homo sapiens (human)											
Homo sapiens											
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.											
1 (bases 1 to 1201)											
L1,W.B., Gruber,C., Jesse,J., and Polayes,D.											
Full-length cDNA libraries and normalization											
Unpublished (2001)											
Contact: Genoscope											
Genoscope - Centre National de Sequencage											
Bp 191 91006 EURY cedex - France											
Email: seqre@genoscope.cns.fr, web : www.genoscope.cns.fr											
Library was constructed by Life Technologies, a division of											
Invitrogen. Contact : Feng Liang Email : fliang@life tech.com URL :											
http://fulllength.invitrogen.com/Invitrogen Corporation 1600											
Faraday Avenue Genoscope sequence ID : CSOD1072CC03NP1.											
Location/Qualifiers											
1..1201											
/organism="Homo sapiens"											
/mol_type="mRNA"											
/db_xref="taxon:9606"											
/clone="CSOD1072YF05"											
/issue_type="PLACENTA COT 25-NORMALIZED"											
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"											
/note="1st strand cDNA was primed with a NcOI-oligo (dT)											
primer. Five prime end enriched, double-strand cDNA was											
digested with NcOI and cloned into the NcOI and EcoR V											
sites of the pCMVSPORT 6 vector. Library was normalized."											
ORIGIN											
Query Match	5.2%	Score 39.2	DB 13	Length 1201							
Best Local Similarity	6.3%	Pred. No. 13									
Matches	34	Conservative 210	Mismatches 296	Indels 0	Gaps 0						
59	TGGAACCGATGATGAAGTATCGCGGCTACGAATGCGCGAAGTCTAAAACTCC	118									
620	KKCGMKKMMKNN	679									
119	TTGCTGCAGTTTGAGCATGTACCCGATGTGATTCGTTGCTCTTGCTATTCAAT	178									
680	KKKKKKNN	739									
179	ACGGCTACGATGTGTGGCGCAGCCCACTGTTTCGGGCGTATTTGCGGGCTCACCG	238									
740	MMANNCKMKKNNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	799									
229	AAATGCTGTCATGCCCTCGTTGTGGCGGAGCGCCCTGGGCGCATCGGCTACCA	298									
800	MCCKKKKKKMKKGMKCMCKMMKKKKCMCKMKMMCKMKMMCKMKMMCKMKMM	859									
299	CATTGCTGTGAATCTCGCGCAGATTCATGCGTTTCAATCCCGCTGATGTGCTCA										

Db	1100	SCMMBSGKCCCKTKMBKCMCMQMBQBMBXKVYVKKKKKKKKKKKKKKKKKKKK	1155
Qy	599	GC 600	
Db	1160	BC 1161	
<hr/>			
RESULT 6			
CNS0091P			
LOCUS			
DEFINITION	CNS0091P	925 bp	DNA linear GSS 03-JUN-1998
ACCESSION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
VERSION	BACR19D16 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
KEYWORDS	AL053013		
SOURCE	AL053013.1 GI:4934461		
ORGANISM	GSS.		
REFERENCE	Drosophila melanogaster (fruit fly)		
AUTHORS	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;		
JOURNAL	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
	Ephydroidea; Drosophilidae; Drosophila.		
	1 (bases 1 to 925)		
	Genoscope.		
	Direct Submission		
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequenage :		
	Bp 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
	Determination of this BAC-end sequence was carried out as part of a		
	collaboration with the Berkeley Drosophila Genome Project (BDGP).		
	The BDGP is constructing a physical map of the Drosophila		
	melanogaster genome using these BACS. For further information		
	please see http://www.fruitfly.org The BDGP Drosophila		
	melanogaster BAC library was prepared by Kazutoyo Oosagawa and		
	Aaron Mammoser in Peter de Jong's laboratory in the Department of		
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,		
	NY. The library is named RPCL-98 and was constructed by partial		
	EcoRI digestion of Drosophila DNA provided by the BDGP from the		
	isogenic strain Y2; cn bw sp, the same strain used for the BDGP's		
	p1 and EST libraries. A more detailed description of the library		
	and how to order individual BAC clones, the entire library, or		
	filters for hybridization from the BACPAC Resource Center can be		
	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.		
FEATURES			
source	Location/Qualifiers		
	1..925		
	/organism="Drosophila melanogaster"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:7277"		
	/clone="BACR19D16"		
	/clone_1fb="RPCL-98"		
	/note="end : TET3"		
ORIGIN			
Query Match	5.2%; Score 39; DB 29; Length 925;		
Best Local Similarity	14.2%; Pred. No. 19;		
Matches	37; Conservative 117; Mismatch 106; Indels 0; Gaps 0		
Qy	195	GGCAGCGCCACGATTTCGCCGCTCATTTTTCGGGGGCTCACCGAATGTGATCAATCC	254
Db	638	SSSSSSSTSSSSSTSSSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	697
Qy	255	CTCTGTTGGGCGCGACGCCCTTGCGCCCATTCGCGCTACCACATTCTGTGAATT	314
Db	698	SSSSSSSTSSSSSVSSSGSKSSTBSGSGBSSSSSSSSSSSSSSSBSTSSSSSSSSST	757
Qy	315	CGCCGACGATTTCTAATCGTTTCATTCGCCGCGCATGCGCAAAAACCCATTGCCG	374
Db	758	CGCTCCCCSYSS	817
Qy	375	TTCATTAATGGATTTCGCGCTTATGCAGAAAGCTTAGCATGCTCGGCGCAGCGCCG	434
Db	818	YMECYTSTSSCGSSSSSSGKGVTGKCGCGGSSSTNGMBHTSSACSSSSSSCGSSSVSS	877

[illegible]

CB657749 835 bp mRNA linear EST 09-APR-2003
 LOCUS OSJNEC1F02.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION clone OSJNEC1F02 5', mRNA sequence.
 ACCESSION CB657749
 VERSION CB657749.1 GI:29661474
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoideae; Oryza.
 1 (bases 1 to 835)
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)
 CONTACT: Rod Wing
 JOURNAL Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 13 row: F column: 02
 Seq primer: gta aac cga cgg cca gtc.
 Location/Qualifiers
 1..835
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEC1F02"
 /cvsue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEC"
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
 XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN
 Query Match 5.0%; Score 38; DB 14; Length 835;
 Best Local Similarity 51.1%; Pred. No. 34;
 Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 252 CGGCTGCTGTGAGGCGAGCGCCCTGAGCGGCATGCGCTACCCACATTGCTGATGA 311
 79 CGACCGGCTCTCGAGCGCGCGCTGCTCGGCGCTTCGCGCGCTGCGGAGATGCCCTC 138
 312 CTTCGCCGACGATTCTATGCGCTTTCATTCCTCGGCTGATGATGATTAACCCATTTC 371
 139 ATCCCGCGGGTGTTCGACCGATGCGGACCCCGACCTGCGCGCTGGAACGCGTGTCT 198
 372 CCGTTCTATTCGGATTTTCGCGCTTATCGACGAGCTTACGACGATCACTGCGGC 425
 199 CTCGGCTACGCGCGCTTCTGTCGACGAGACGTGGCTGCGGACCTCTGCGGC 252

RESULT 9
 BH887860 697 bp DNA linear GSS 07-AUG-2002
 LOCUS LB01712a.d.SP6.1 Leishmania major Friedlin BAC Library Leishmania
 DEFINITION major genomic clone LB01712a, genomic survey sequence.
 ACCESSION BH887860
 VERSION BH887860.1 GI:22133383
 KEYWORDS GSS.
 SOURCE Leishmania major
 ORGANISM Leishmania major

Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
 Leishmania.
 1 (bases 1 to 697)
 Myler,P.J., Vogt,C., Munden,H., Robertson,L., Slek,E.,
 Ravelin,A.G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E.,
 Stuart,K. and Ragland,M.
 Leishmania major Friedlin BAC End Sequences
 Unpublished (2002)
 Other GSSs: LB01712a.d.T7.1
 CONTACT: Myler PJ
 Seattle Biomedical Research Institute
 4 Nickerson Street, Seattle, WA 98109-1651, USA
 Tel: 206 284-8846
 Fax: 206 284-0313
 Email: mylerpj@bri.org
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..697
 /organism="Leishmania major"
 /mol_type="genomic DNA"
 /strain="Friedlin"
 /db_xref="taxon:5664"
 /clone="LB01712a"
 /lab_host="E. coli GeneHog + TrfA"
 /clone_lib="Leishmania major Friedlin BAC library"
 /note="Vector: pCG270; Site 1: HindIII; Genomic DNA from
 Leishmania major Friedlin in agarose blocks was partially
 digested with HindIII, size selected, and ligated with
 HindIII-digested pCG270 vector DNA. 10368 clones were
 picked and arrayed in 384- and 96-well plates. Library
 construction and arraying was carried out by Resgen
 Corporation and clones and filters are available from
 them"

ORIGIN
 Query Match 5.0%; Score 37.4; DB 28; Length 697;
 Best Local Similarity 52.2%; Pred. No. 46;
 Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 181 GCTACGAAATGCTGGGAGACCCCACTGTTTCGCGCTGATTTTCGGGCTCAACGAA 240
 579 GCGAGCGATGCTGCTGTGCGGACTCTCCACCCCAACGATCTCGCGCGCGCG 520
 241 ATGCTGCTATCGGCTGCTGTGAGGCGAGCGCCCTGAGCGGCATCGGCTCACCA 300
 519 CACGAGTCGGCGCGCTGCTGCGAGCCTCGTCCGCTTTGCGGCAACCGTTACCAAG 460
 301 TTGCTGTGAACCTTCGCGACGATTTCTATGCGCTTTCA 339
 459 CCACTGTCACTGTGATCATGCTCTCGTGGCCATTCGA 421

RESULT 10
 BX382084/c 1108 bp mRNA linear EST 08-MAY-2003
 LOCUS BX382084 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1074YC24 5-PRIME, mRNA sequence.
 ACCESSION BX382084
 VERSION BX382084.1 GI:30455015
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1108)
 Li,W.B., Gruber,C., Jesse,J., and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

us-10-608-504-2.rst

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1074BB12QPI.

source

1. .1108

```

/clisue_type="PLACENTA.COT.25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA.COT.25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime and enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match	5.0%;	Score 37.4;	DB 13;	Length 1108;
Best Local Similarity	16.3%;	Pred. No. 57;		
Matches	65;	Conservative 145;	Mismatches 189;	Indels 0;
				Gaps 0

[illegible]

LOCUS	EG845919	730 bp	mRNA	linear	EST 29-MAY-2001
DEFINITION	1024011F06.x1	C. reinhardtii CC-1690, normalized,	lambda Zap II		
	Chlamydomonas reinhardtii cDNA, mRNA sequence.				

ACCESSION	BG845919	GI:14227103
VERSION	BG845919.1	
KEYWORDS	EST.	
SOURCE	Chlamydomonas reinhardtii	

ORGANISM

REFERENCE	1 (bases 1 to 730)
AUTHORS	Grossman,A., Davies,J., Federspiel,N., Harris,E., Lefebvre,P.,

TITLE

JOURNAL
COMMENT
Unpublished (2000)
Contact: Charles Hauser

Duke University

Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
Location/Qualif

Source

1. .730 "Chl" = "chlorophyll"

```

/cjOne.lib="C:\reinhardt\CC-1690, normalized, Lambda Zap
II"
/note="Vector: Bluescript II SK-, Site 1: EcoRI, Site 2:
XhoI; This library, constructed by John Davies and Jeffrey
Mcdermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP (acetate-containing) medium in the
light. TAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
POLYA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
Bluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exsistit (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."

```

Query Match	4.9%	Score 37.2;	DB 12;	Length 730;
Best Local Similarity	49.5%;	Pred. No. 53;		
Matches 96;	Conservative 0;	Mismatches 98;	Indels 0;	Gaps 0

QY 202 CCACGTTTCCGGCTATTTTGGCGGGCTCAACCGAAATGCTGTATCGCCCTCGTT 261

Db 611 CCCCGETTCAAGGGCGGCAAGCCGGCCCCCCCCCAAGATTCAAGGCATCGGCGCCCGGC 632

QY 262 GTGGCGGAGCGCCCCCTGGGCGCCATCGCGCTCACACATTGCTGTAACTTCCGCCAC 321

Db 611 TTGTGGCGCGGGTCTGACCCGCCCTGATCAGCGAGGTGTCACAGTGTCCAGTAGC 572

QY 322 GTATTCTATGGCTTTTCATTCCCGCTGATGTGTGTCAAATACCCCATTTGCCCGTTCTAT 381

Db 571 GACGCCATCGACATAGCTCCCCGCTGGCGCGCTGAGAGAGGGCTCTATGTTGGGCATCAGC 512

QY 382 TCGGTTTTCGCGCT 395

Db 511 TCGGCGCCCGCGGT 498

LOCUS	154 bp	EST 19-OCT-2000
DEFINITION	BR101158 mRNA	linear
	601734733F1 NCI_CGAP_Mam1 Mus musculus	cdna clone
	mRNA sequence.	
	IMAGE:3983528	5'

ACCESSION	BF101158
VERSION	BF101158.1
KEYWORDS	GI:10883684
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
PROJECT	
REMARKS	

ORGANISM

REFERENCE
Mammals, Eutheria, Rodentia, Sciuromorphi: Muridae, Mus.
1 (bases 1 to 1454)
NH-MGC <http://mgc.ncl.nih.gov/>.

JOURNAL

COMMENT:

<http://image.11n1.gov>

Plate: L1AM9184 row: b column: 09
 High quality sequence stop: 606.
 Location/Qualifiers

FEATURES

source

1..1454
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3983528"
 /issue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match
 Best Local Similarity 50.0%; Pred. No. 74;
 Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
 114 CTCCTCTGTCGAGTTTGGGCAATGATCCGATGGTATGGTGTCTCTGTTAT 173
 1159 CTCCTCTGTCGAGTTTGGGCAATGATCCGATGGTATGGTGTCTCTGTTAT 1100
 174 TCATATGGGCTAGAGATGGGAGAGCCGACATGTTTCGAGGCTATTTGCGGGCTC 233
 1099 TGGCTGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1040
 234 CACCGAATGCTGTCATCGCCCTGTTGTGGGCGAGCGCCCTGGGCGCATGGCGCT 293
 1039 CGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
 294 CACGAC 299
 979 CTCCTC 974

RESULT 13
 BH805539/c 464 bp DNA linear GSS 25-APR-2002
 LOCUS 1008061A12.x1 1008 - Rescued Grid 1 Zea mays genomic, genomic
 DEFINITION survey sequence.

ACCESSION BH805539
 VERSION BH805539.1 GI:20322930

KEYWORDS

SOURCE

ORGANISM

Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 464)

REFERENCE
 AUTHORS Maize genomic sequences found using engineered Rescued transposon
 TITLE Unpublished (2001)
 JOURNAL Contact: Walbot V
 COMMENT Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1008061 row: 27
 Class: transposon-tagged.
 Location/Qualifiers

FEATURES

source

1..464
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="mixed background W23/A18/B73"
 /db_xref="taxon:4577"
 /tissue_type="leaf"

ORIGIN

Query Match
 Best Local Similarity 52.4%; Pred. No. 55;
 Matches 108; Conservative 0; Mismatches 92; Indels 6; Gaps 1;
 223 TTGCGGGCTCCACCAATGCTGTCATCGCCCTGTTGTGGCGAGCGCCCTGGGC 282
 278 TTGCGGACTCCACGAGCGTGGGCTCATGCTCTGGCGCGCGCGGCGGCGCAAC 219
 283 GCGATCGGCTCACCACATGCTGTAATTCGCGCACTATTTATGCTTATTC 342
 218 GCGCGCGAGTCCCTTCATGACCGTCTTTCGCACTCATGAGCGCTTGGCGGC 159
 343 CGGCTG-----CATGTCGTAATAACCCCATTCCTGTTTATTCGTTTGGCGCTT 396
 158 GCGCTCAGCATCCAGCATGCTGTCACACCGGTGTCATGCTCTTCGATTTGCTAC 99
 397 ATGCAAGAGCTTACGACGATCTGC 422
 98 CTCGCGATCGGCTCGCGCGTGGCTTC 73

RESULT 14
 CDS0168R 1201 bp DNA linear GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACN15K14 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION AL106545
 VERSION AL106545.1 GI:5622456

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)
 GSS.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1201)

REFERENCE
 AUTHORS Direct Submission
 TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage;
 JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage;
 COMMENT BP 191 91006 EVRY cedex - FRANCE (E-mail: seque@genoscope.cns.fr)
 - Web: www.genoscope.cns.fr
 determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (BDGP) -
 http://www.edgp.sbi.ac.uk. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CNRS (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES

source

1..1201
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN15K14"
 /clone_lib="DrosBAC"

Tue Apr 20 06:47:17 2004

us-10-608-504-2.rst

Page 8

ORIGIN /plasmid="pBelobAC11"
/note="end : T7"

Query Match 4.9%; Score 36.8; DB 29; Length 1201;
Best Local Similarity 20.4%; Pred. No. 87;
Matches 73; Conservative 131; Mismatches 153; Indels 1; Gaps 1;

Qy 38 TGTGGCCATCCAGAGCCGCTGAAACAGATGATTAAGTTATCGCCGCTACGAATCG 97
Db 823 TTTCCCMAMCMOMAMMMCCCTTCMAGTANCTCMAMMMMGSGGCAAMVVMV 882
Qy 98 CGCAAGGTCTAAAACCTCCCTGCTGAGGTTGGGCAATGACCCGATTTGATTCGT 157
Db 883 VVMVMSYCCCMAGGACCHTCAATGACGAGTATGATGATGATGATGATGATGAT 942
Qy 158 TTGATCTCTGTTATTCATACGAGCTACGATGATGATGATGATGATGATGATGAT 216
Db 943 YTGAGACCCGCTTTMWSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSSGSS 1002
Qy 217 CTGATTTTCGGGGCTCCACGAAAGCTGCTGATGCTGCTGCTGCTGCTGCTGCT 276
Db 1003 SSSCBSTGTCTYSS 1062
Qy 277 CTGGGGCCGACGCTCACACATGCTGATGATGATGATGATGATGATGATGATGAT 336
Db 1063 CBCCCCGCKCCGSGSTSKSTGKTGKSSGSGGSGSSSSSSSSSSSSSSSSSSSS 1122
Qy 337 TCATTCGCGCTGATGTGCAAAACCCATGCGCCGTTCTATTCGCTGCTGCTGCT 394
Db 1123 TTTTBSSSSTSTSTSSGSBRTBSSBSSTSTSTSTSTSTSTSTSTSTSTSTSS 1180

RESULT 15
BE013856/c 539 bp mRNA linear EST 09-JUL-2000
LOCUS 125354 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BE013856
ACCESSION BE013856
VERSION BE013856.1 GI:8274843
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 539)
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -timematch 12 options.
PCR primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCACAGTACAGAG
Plate: 52 row: G column: 21
Seq primer: ATTGAGTACACTATAG.
Location/Qualifiers
1..539
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"

ORIGIN /lab_host="DH10B"
/clone_id="MARC 1PIG"
/note="Vector: PCMV SPOR6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

Query Match 4.9%; Score 36.6; DB 10; Length 539;
Best Local Similarity 54.0%; Pred. No. 67;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 211 TCCGCGCTGATTTTCGGGGGCTCCACCGAATGCTGCTCATCGCCCTGTTGTGGCGCA 270
Db 445 TCTGCTGGGCTTCCGAGGCGCCAGGCTGAGGGGTGTGAGGGCTGTGACAGATCA 386
Qy 271 GCGCCCTGGGCGGCTACGCTCAACATGCTGTGTAATTCCGCCACGTAATTCTAT 330
Db 385 CCGGCCAGGGAAGCAGCACCCGCCACCTCTCTGACAGCCAGGAGCGGAGTCAAG 326
Qy 331 GCGTTTATTCGCGCTGC 349
Db 325 GTGTCTGACGCGCTGC 307

Search completed: April 17, 2004, 20:01:43
Job time : 2440.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:12 ; Search time 70.6156 Seconds
(without alignments)
1004.303 Million cell updates/sec

Title: US-10-608-504-3

Sequence: 1 MORTDERHSHLSLEVPSPKAL.....LIFLGLTLTRYFLGKAAK 251

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	4 AAG90035	Aag90035 C glutam
2	1265	99.7	251	4 AAB86247	Aab86247 C. glutam
3	428.5	33.8	237	4 AAG31154	Aag31154 C glutam
4	214	16.9	338	5 ABP66261	Abp66261 Bifidobac
5	197.5	15.6	261	6 ABM69066	Abm69066 Photobac
6	197	15.5	245	5 ABP57756	Abp57756 E. coli L
7	190	15.0	266	6 ADA35238	Ada35238 Actinobac
8	189.5	14.9	263	6 ADA33199	Ada33199 Actinobac
9	185.5	14.6	255	7 ADC95293	Adc95293 E. faeciu
10	177.5	14.0	218	6 ABU00481	Abu00481 S. pneumo
11	177.5	14.0	218	6 ABP81300	Abp81300 Streptoco
12	174.5	13.8	235	5 ABM54958	Abm54958 Lactococ
13	169	13.3	235	5 ABM47959	Abm47959 Listeria
14	167	13.2	230	5 ABP30530	Abp30530 Streptoco
15	167	13.2	231	5 ABP28562	Abp28562 Streptoco
16	161.5	12.7	242	5 ABP40309	Abp40309 Staphyloc
17	158.5	12.5	230	4 AAG81587	Aag81587 S. epider
18	150.5	11.9	226	6 ADB08660	Adb08660 Allotococ
19	148.5	11.7	214	6 ADB08658	Adb08658 Allotococ
20	148	11.7	228	6 ABM72815	Abm72815 Staphyloc
21	141.5	11.2	247	6 ADA34737	Ada34737 Actinobac
22	114	9.0	475	4 AAU64076	Aau64076 Propionib
23	114	9.0	475	6 ABM60595	Abm60595 Propionib
24	112.5	8.9	876	5 ABP55529	Abp55529 Bifidobac
25	112	8.8	379	5 ABB55268	Abb55268 Lactococ

26	111	8.7	113	3 AAY74753	Aay74753 Neisseria
27	105	8.3	417	7 ADD46525	Add46525 Human pro
28	105	8.3	417	7 ADD46529	Add46529 Human pro
29	104	8.2	377	6 ADB12483	Adb12483 Allotococ
30	104	8.2	414	6 ADB12481	Adb12481 Allotococ
31	104	8.2	442	6 ADB12479	Adb12479 Allotococ
32	102	8.0	112	3 AAY74752	Aay74752 Neisseria
33	101.5	8.0	746	7 ADB75242	Adb75242 Prostata
34	101.5	8.0	766	4 AAM79259	Aam79259 Human pro
35	101.5	8.0	816	6 ADA5075	Ada5075 Human pro
36	101.5	8.0	820	5 ABP56272	Abp56272 Human TCH
37	101.5	8.0	1203	4 ABG05471	Abg05471 Novel hum
38	101.5	8.0	1597	4 ABG19902	Abg19902 Novel hum
39	101.5	8.0	1597	4 ABG10812	Abg10812 Novel hum
40	101.5	8.0	1597	4 ABG10253	Abg10253 Novel hum
41	101.5	8.0	1597	4 ABG03989	Abg03989 Novel hum
42	101.5	8.0	1599	4 ABG07884	Abg07884 Novel hum
43	101.5	8.0	1599	4 ABG14450	Abg14450 Novel hum
44	101.5	8.0	1784	4 ABG09148	Abg09148 Novel hum
45	98.5	7.8	478	3 AAG15416	Aag15416 Arabidops

ALIGNMENTS

RESULT 1
AAG90035
ID AAG90035 standard; protein, 251 AA.
XX
AC AAG90035;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 3789.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PP 18-DEC-2000; 2000EP-00127668.
XX
PR 16-DEC-1999; 99JP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX
PA (KJOW) KJOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX N-PSDB; AAH65254.
XX
PT Novel polynucleotides derived from Corynebacterium, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 3789; 245pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of Corynebacterium, measuring expression amount and analyzing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium, and identifying a homologue of a gene derived from
XX Corynebacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office

XX Sequence 251 AA;

Query Match 100.0%; Score 1269; DB 4; Length 251;
 Best Local Similarity 100.0%; Pred. No. 3.3e-137;
 Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREYIAQGLKTSIAAGIGMYPGIAFGLVIOY 60
 DB 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREYIAQGLKTSIAAGIGMYPGIAFGLVIOY 60
 QY 61 GYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVFRHVFYAFSPPLHYVK 120
 DB 61 GYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVFRHVFYAFSPPLHYVK 120
 QY 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
 DB 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
 QY 181 IKGLEPALSFLVTLTLDSCRTKQIPSLLAGSFTTALVTPGOALFAALLIFLGILT 240
 DB 181 IKGLEPALSFLVTLTLDSCRTKQIPSLLAGSFTTALVTPGOALFAALLIFLGILT 240
 QY 241 IRYFFLGKAAK 251
 DB 241 IRYFFLGKAAK 251

RESULT 2

ID AAB86247 standard; protein; 251 AA.

AC AAB86247;

DT 05-SEP-2001 (first entry)

DE C. glutamicum bnf protein.

KM L-aminic acid production; bnf; bnf; branched-chain amino acid;

KM Coryneform bacterium; leucine; isoleucine; valine; medicine;

XX animal nutrition.

OS Corynebacterium glutamicum.

PN EP1096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

PR 27-OCT-1999; 99DB-01051708.

PA (DEGS) DEGUSA AG.

PA (KERU) FORSCHUNGSZENTRUM JUELICH GMBH.

PI Kernerkecht N, Eggeling L, Sahn H, Pfeffertje W,

DR WPI; 2001-391595/42.

DR N-PEDB; AAH21109, AAH21110.

XX New export genes from coryneform bacteria, useful for increasing

PT fermentative production of branched-chain amino acids.

PS Claim 1a; Page 15-16; 23pp; German.

CC This invention describes a novel isolated polynucleotide (i) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least

CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (i), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC bnf and/or bnf genes (or equivalent sequences) are amplified.
 CC especially overexpressed; and (d) method for isolating the bnf and/or
 CC bnf genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids;
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence represents
 CC the Corynebacterium glutamicum ATCC 14752 bnf protein described in the
 CC method of the invention

XX Sequence 251 AA;

Query Match 99.7%; Score 1265; DB 4; Length 251;
 Best Local Similarity 99.6%; Pred. No. 9.4e-137;
 Matches 250; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREYIAQGLKTSIAAGIGMYPGIAFGLVIOY 60
 DB 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREYIAQGLKTSIAAGIGMYPGIAFGLVIOY 60
 QY 61 GYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVFRHVFYAFSPPLHYVK 120
 DB 61 GYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVFRHVFYAFSPPLHYVK 120
 QY 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
 DB 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
 QY 181 IKGLEPALSFLVTLTLDSCRTKQIPSLLAGSFTTALVTPGOALFAALLIFLGILT 240
 DB 181 IKGLEPALSFLVTLTLDSCRTKQIPSLLAGSFTTALVTPGOALFAALLIFLGILT 240
 QY 241 IRYFFLGKAAK 251
 DB 241 IRYFFLGKAAK 251

RESULT 3

ID AAG93154 standard; protein; 237 AA.

AC AAG93154;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6908.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

XX Corynebacterium glutamicum.

PN EP108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI, 2001-376931/40.
 DR N-PSDB; AAH68373.
 XX Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 17; SEQ ID NO 6908; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC *Corynebacterium* *Corynebacterium*, and identifying a homologous of a gene derived from
 CC *Corynebacterium* *Corynebacterium*. *Corynebacterium* bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 CC
 SQ Sequence 237 AA;
 XX
 Query Match 33.8%; Score 428.5; DB 4; Length 237;
 Best Local Similarity 40.1%; Pred. No. 1.5e-40;
 Matches 87; Conservative 46; Mismatches 81; Indels 3; Gaps 3;
 XX
 QY 31 ETGQGLTSLAGLGMVPIPIGIAFGLVITQGYEWMAAPFSGIIPFGSTEMLVIALV-VG 89
 DB 8 EIRGGIGETLVGLGIPGLAFLGLMVGTFAMMTPIFSVIVYSGNSEFLAIGVAVG 67
 QY 90 AAPLGAIALTTLLNPRHVFYAFSPFLHYVKNPIARFYGVFALIDEAVVTARAPAG-WS 148
 DB 68 IGFESA-AVAGFVFNRRHIFGILTFPRHRKIGAGRAYTVALTDESVAIVGARPPGDS 126
 QY 149 AMRLIMQIAFHSYVYFGILTVGVAIAELIPPEIKGLEFALCSFTVTLTDSCTKQIPSL 208
 DB 127 GTRVLTVQILCOALMTVPIGILGALVGVFPDDIKGMDFAITLALFVLAWEAKNNKDYSL 186
 QY 209 LLLAGSFTIALVVIPOALFALLIFLGLITRYPF 245
 DB 187 PLFAVVALVSGFVAPEQMLVITLTYFPLILLRVF 223
 XX
 RESULT 4
 ID ABP66261 standard; protein; 338 AA.
 XX
 AC ABP66261;
 XX
 DT 19-NOV-2002 (first entry)
 XX
 DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:1005.
 XX
 KM Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
 KM antiarrhythmic; antibacterial; inhibitor of *Salmonella*; detection;
 KM identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
 KM rotavirus; food composition; pharmaceutical composition.
 XX
 OS Bifidobacterium longum.
 XX
 PN EP1227152-A1.
 XX
 PD 31-JUL-2002.
 XX
 PF 30-JAN-2001; 2001EP-00102050.
 XX
 PR 30-JAN-2001; 2001EP-00102050.
 XX
 PA (NEST) SOC PROD NESTLE SA.
 XX
 DR WPI, 2002-668397/72.

XX
 PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
 PT a probe or primer for detecting and/or identifying Bifidobacterium longum
 PT in a biological sample.
 XX
 PS Claim 3; SEQ ID NO 1005; 80bp; English.
 XX
 CC The present invention describes a polynucleotide (I) comprising a
 CC sequence of a Bifidobacterium genome selected from the nucleotide
 CC sequences given in AB081842 and AB081843, or a sequence exhibiting at
 CC least 90% identity or which hybridises with the sequences given in
 CC AB081842 and AB081843. Also described is a polynucleotide (II) encoding a
 CC fusion protein, comprising a sequence selected from 1097 sequences given
 CC in ABP65288 to ABP6335, ligated in frame to a polynucleotide encoding a
 CC heterologous polypeptide. (II) has antiarrhythmic and antibacterial
 CC activities, and can be used as an inhibitor of *Salmonella*. (I) (which is
 CC a probe) is useful for the detection and/or identification of
 CC Bifidobacterium longum in a biological sample. A carrier containing the
 CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
 CC used for preventing and/or treating diarrhoea brought about by pathogenic
 CC bacteria and/or rotavirus. The carrier is a food composition selected
 CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
 CC products, ice-creams, fermented cereal based products, milk based
 CC powders, infant formula, pet food or a pharmaceutical composition
 CC selected from tablets, liquid bacterial suspensions, dried oral
 CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
 CC (I) is useful in DNA arrays or chips to carry out analysis of the
 CC expression of the Bifidobacterium genome. AB081844 to AB081850 represent
 CC Bifidobacterium related nucleotide sequences given in the Sequence
 CC Listing from the present invention but not mentioned further within the
 CC specification. N.B. The sequence data for this patent is not represented
 CC in the printed specification but is based on sequence information
 CC supplied by the European Patent Office
 CC
 SQ Sequence 338 AA;
 XX
 Query Match 16.9%; Score 214; DB 5; Length 338;
 Best Local Similarity 29.9%; Pred. No. 1.1e-15;
 Matches 64; Conservative 37; Mismatches 85; Indels 28; Gaps 8;
 XX
 QY 44 LGMPDIGAPGLVITQGYEWMAAPFSGIIPFGSTEMLVIALV-VG-APIGLAIALTTLL 102
 DB 99 LGPLFLGASVGLINGTGFSFVWPMCSAPFAGSMFVYVNLISLFPNLAGE-LIALM 157
 QY 103 VNFRRVFYAFSPFLHYVKNPIARF-----YVFALIDEAVAV--TAARPAGWS-AMR 151
 DB 158 VNARHLFYGISM-----LGKFKGLGWRKRPYLPFCWCBETFAINSTAKIPAGIDRGWF 209
 QY 152 LISMQIAFHSYVYFGILTVGVAIAELIPPEIKGLEFALCSFTVTLTDSCTKQIPSL-L 210
 DB 210 YFWVTLGNQILVWVGATLGLIGHLPNTDGLDFVLTALFLVFLDQWMDGRKRRRLSA 269
 QY 211 LAGSFTIALVVIPOALFALLIFLGLITRYPF 245
 DB 270 VIQVLTSLACLLIFGANDFMIPIAMILLIFVAL 303
 XX
 RESULT 5
 ID ABM69066 standard; protein; 261 AA.
 XX
 AC ABM69066;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens protein sequence #2163.
 XX
 KM Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KM detection; food; gene expression; plant; animal; microorganism; toxin;
 KM antibiotic; bioplastic; virulence factor; disease model; plague;
 KM whooping cough.
 XX
 OS Photorhabdus luminescens.

AC ADA35238;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #2399.
 XX
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 XX
 DR N-PSDB; ADA31112.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 XX
 PT for diagnosing a bacterial disease, as components of antibacterial
 XX
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 XX
 PT plants.
 XX
 PS Example; SEQ ID NO 6525; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 XX
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 XX
 CC for diagnosing a bacterial disease, as components of antibacterial
 XX
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 XX
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 XX
 CC compounds for the ability to interfere with the A. baumannii life cycle
 XX
 CC or to inhibit A. baumannii infection, and as biocontrol agents for an A.
 XX
 CC baumannii protein.
 XX
 SO Sequence 266 AA;
 XX
 Query Match 15.0%; Score 190; DB 6; Length 266;
 Best Local Similarity 26.6%; Pred. No. 4.6e-13;
 Matches 62; Conservative 43; Mismatches 98; Indels 30; Gaps 7;
 QY 34 QGLKTSIAAGGWPPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVVGAAPL 93
 DB 36 QGANDSQRIYVTVTFVGFARFVSASQFTFMEAFPLSCSWYAGASQFLVVALLAGSSI 95
 QY 94 GAIALTLLVNFHFVFA-----FSPFLAVVKNPIARFYSVPLIDBAVA-----VTAAR 143
 DB 96 WLTAITVIALDIRHVLGPALYNLIPKLNKKTAV-----WAWGLTBVVASGMITQLSOR 151
 QY 144 PAGSAMRLTISMOIAFYSWY-----VFGGLTGVAIAELIPFEIKGLFALCSLFTVTLTLD 199
 DB 152 ROOMSESMGLSLFSWMSWASGSLGLFADQVAHLPKFLQALDLPRLPLFISFLAA 211
 QY 200 CRICKQIPLSLLAGISFTIALVTPGQALFPAALIFLGLT-----IRIFFL 246
 DB 212 FERKHS-----LVVAVSLVSAALACYWINTLSAA--IFGIISGLIAGLGFXYVL 258

XX
 DE Acinetobacter baumannii protein #360.
 XX
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 XX
 DR N-PSDB; ADA29073.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 XX
 PT for diagnosing a bacterial disease, as components of antibacterial
 XX
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 XX
 PT plants.
 XX
 PS Example; SEQ ID NO 4486; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 XX
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 XX
 CC for diagnosing a bacterial disease, as components of antibacterial
 XX
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 XX
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 XX
 CC compounds for the ability to interfere with the A. baumannii life cycle
 XX
 CC or to inhibit A. baumannii infection, and as biocontrol agents for an A.
 XX
 CC baumannii protein.
 XX
 SO Sequence 263 AA;
 XX
 Query Match 14.9%; Score 189.5; DB 6; Length 263;
 Best Local Similarity 24.0%; Pred. No. 5.2e-13;
 Matches 49; Conservative 54; Mismatches 96; Indels 5; Gaps 5;
 QY 40 LAAAGLWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVVGAAPLGAIALT 99
 DB 53 LPLSIVIPWALIASMAVHGLSFKYKALMSGIVFAGAAQVLSVMWEGASLITIVT 112
 QY 100 TLLVNFHFVFAFSPFLAVVKNPIARFYSV-PALIDEXYATTAAPACMSAMRLTSMOIA 158
 DB 113 IFFLTAQHFTVIALTRNDISILPLSKRTLGLLDELFAVSVEKKGPOI-LEGAGIC 171
 QY 159 FHSWVFGGLTGVAIAELIPFEIK-GLFALCSLFTVTLTLDSCRTKQIPLSLLAGLS-F 216
 DB 172 FIDFWVFSVLGILATLALPILNLNTHDPSITAIIVAMIVPCCKKPPWAGILMTCVSGF 231
 QY 217 TIALVTPGQALFPAAL-IFGLL 239
 DB 232 VLKFEHIEGAILISGLHGFIVAI 255

RESULT 9
 ADA35238
 ID ADC95293 standard; protein; 255 AA.
 XX
 AC ADC95293;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 4920.
 XX

KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KW abdominal-pelvic infection.
OS Enterococcus faecium.
XX US6583275-B1.
XX
XX 24-JUN-2003.
XX
XX 30-JUN-1998; 98US-00107532.
XX
XX 02-JUN-1997; 97US-0051571P.
XX 14-MAY-1998; 98US-0085586P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucetbe-Stamm LA, Bush D;
XX
XX MPI; 2003-799836/75.
XX
XX N-PSDB; ADC91639.
XX
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
XX Enterococcus faecium polypeptide useful for detection, prevention and
XX treatment of a pathological condition resulting from a bacterial
XX infection.
XX
XX Example 1; SEQ ID NO 4920; 243pp; English.
XX
XX The invention relates to an isolated nucleic acid derived from
XX Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX one of 10 fully defined sequences given in the (or comprising 40
XX sequential nucleotides chosen from any of the nucleic acids, its
XX complement or sequences hybridising to it). Also included are a
XX recombinant vector comprising the nucleic acid operably linked to
XX transcription regulatory element, a cell comprising the vector and a
XX single-stranded probe comprising the nucleic acid. The nucleic acids are
XX chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX The nucleic acid is useful for diagnosing pathological conditions
XX resulting from E. faecium bacterial infection (e.g. urinary tract
XX infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX infection) and for screening drugs such as agonists and antagonists. The
XX nucleic acid is useful for recombinant production of Candida albicans -
XX derived peptides or antisense polypeptides. Pharmaceutical compositions
XX and vaccines containing the nucleic acid are useful for preventing or
XX treating Enterococcus faecium infections. The present sequence represents
XX one if the disclosed E. faecium proteins.
XX
XX Sequence 255 AA;
SQ
Query Match 14.6%; Score 185.5; DB 7; Length 255;
Best Local Similarity 25.0%; Pred. No. 1.4e-13;
Matches 63; Conservative 51; Mismatches 105; Indels 33; Gaps 10;
16 SKAALEPPDKGRRREIKQGLTKSLAGLGNYPFGIGLVLVQYGVEMAAALPFGULF 75
15 AKGVALLDER---LDIKTKIKDTLPVFGYTGIGLFGVGRAGGHPHVVTLMSLVY 70
76 AGSTEMVIALVGAAPGATATLTLVNFHVFYAFSPFLHVVKNPIAFYGVFALI-D 134
71 AGSAQFIVSMASHPSLISVSTFLVNSRMIMSWTIAFYFKNNLLQNLITGLTMD 130
135 EAYAVTARPAQWMSMRISMOIAFH-----SY--WYFGGTLGYAIALPFAIK-G 183
131 ESFAL-----GNNALNTTGGKLNFRMNTNMISYITWGGSLVGLLGNFTITDPKFG 184
184 LEFALCSLFTVLT---TDSCKRYKQIPLSLAGLSFT---IALVPIGQALFALLIF- 235
185 LDFALVAMFGLYLVQVSDRNTSKRL-QLITLITLIVVGLIFPSMLIVVVTLLG 243
236 --IGLLTRVYF 245
244 CGLGWVTKHAF 255

RESULT 10
ABU00481
ID ABU00481 standard; protein; 218 AA.
XX
XX AC ABU00481;
XX
XX DT 23-OCT-2003 (revised)
XX DT 11-FEB-2003 (first entry)
XX
XX DE S. pneumoniae type 4 strain protein from coding region #48.
XX
XX KW Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
XX antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX gene therapy; vaccine.
XX
XX OS Streptococcus pneumoniae; type 4 strain.
XX
XX PN WO200277021-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 27-MAR-2002; 2002WO-1B002163.
XX
XX PR 27-MAR-2001; 2001GB-00007658.
XX
XX PA (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX PI Maigman V, Tettein H, Fraser C;
XX
XX DR MPI; 2003-040579/03.
XX
XX DR N-PSDB; ABX05760.
XX
XX PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
XX useful as medicaments for treating or preventing a disease or infection
XX due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
XX ear infection.
XX
XX PS Claim 1; SEQ ID NO 96; 56pp; English.
XX
XX The invention relates to a protein comprising or having at least 50%
XX identity to any of the 2469 amino acid sequences, identified in the
XX specification (available on a computer readable format), or its fragment,
XX expressed from 2469 identified DNA coding regions from the
XX Streptococcus pneumoniae type 4 strain genomic sequence appearing as
XX AB56454. Also included are an antibody which binds one of the proteins,
XX treating a patient by administering the protein, DNA or antibody (in a
XX composition), a kit comprising first and second primers, which are the
XX nucleic acid cited above or fragments between nucleotides 8-100 of a
XX sequence not defined in the specification, for amplifying a target
XX sequence contained within a Streptococcus nucleic acid sequence, where
XX the first primer is substantially complementary to the target sequence
XX and the second primer is substantially complementary to the complement of
XX the target sequence, and where the pairs of the primers having
XX substantial complementarity define the termini of the target sequence to
XX be amplified, assay comprising contacting a test compound with the
XX protein, and determining whether the test compound binds to the protein
XX and a Streptococcus pneumoniae bacterium, where one or more genes
XX encoding the proteins has been rendered inactive. The proteins, nucleic
XX acid molecules, antibody and compositions are useful as medicaments for
XX treating or preventing a disease or infection due to streptococcus
XX bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
XX media or ear infection. They are also useful in developing vaccines,
XX immunomodulators and antibiotics. The methods are useful for identifying
XX immunodominant proteins. The present sequence is one of the 2469 proteins
XX expressed by the sequence data for this patent did not form part of the printed
XX Note: The sequence data for this patent did not form part of the printed
XX at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
XX standardise OS field)
SQ Sequence 218 AA;

[illegible]

CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN604-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins

XX Sequence 230 AA;

Query Match 13.2%; Score 167; DB 5; Length 230;

Best Local Similarity 25.0%; Pred. No. 1.7e-10; Mismatches 97; Indels 34; Gaps 9;

Matches 58; Conservative 43; Mismatches 97; Indels 34; Gaps 9;

QY 24 DKCYRRYEIAQGLTSLAAGLWYPIGAFGLVIOGYEWMAAPLPSGLIFAGSTEMLV 83

DB 2 EKGFK-----EGVKDALPTALGYISIGLAFIVASASDLSAIEVGLMSALVYGSSAQFAM 56

QY 84 IALVGAAPLGAIALTLTLVNFHFVAFSPFLV-----VKNPFAFYSVFLI-D 134

DB 57 CALLAKADLMTITMTVFLVNLRLMS-----LHATTIKSAHLNQLA-----IGTLITD 108

QY 135 EAYAVTAPAPAGWS---AMRLISMQIAFHSYVFGGLTGAVALLELPF-EIKGLEFALC 189

DB 109 ESYGVLLGEALHKKVSPSW-MEGNNVMSYLTWISTIGTLGSTITPNPEMFGDLFALV 167

QY 190 SLFVTLTIDSCRTKQIPSLLAGLSTFIALVTPQALPALLIFIGLTI 241

DB 168 AMFTGLFV-----FOLFGMLSDGKRLVYVIVLASVGLSTFLATPLSGALSV 213

RESULT 15

ABP28562 ID ABP28562 standard; protein; 231 AA.

AC ABP28562;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 6300.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

OS Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB004789.

PR 27-OCT-2000; 2000GB-00026333.

PR 24-NOV-2000; 2000GB-00028727.

PR 07-MAR-2001; 2001GB-00005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Masiagnani V, Margarit Y Rosi, Grandi G, Fraser C;

XX Tettelein H;

XX WPI, 2002-352536/38.

XX DR N-PSDB; ABN69193.

PT New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.

PS Claim 1; Page 3795; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP10895) from group B
 CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN604-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins

SQ Sequence 231 AA;

Query Match 13.2%; Score 167; DB 5; Length 231;

Best Local Similarity 25.0%; Pred. No. 1.7e-10; Mismatches 97; Indels 34; Gaps 9;

Matches 58; Conservative 43; Mismatches 97; Indels 34; Gaps 9;

QY 24 DKCYRRYEIAQGLTSLAAGLWYPIGAFGLVIOGYEWMAAPLPSGLIFAGSTEMLV 83

DB 3 EKGFK-----EGVKDALPTALGYISIGLAFIVASASDLSAIEVGLMSALVYGSSAQFAM 57

QY 84 IALVGAAPLGAIALTLTLVNFHFVAFSPFLV-----VKNPFAFYSVFLI-D 134

DB 58 CALLAKADLMTITMTVFLVNLRLMS-----LHATTIKSAHLNQLA-----IGTLITD 109

QY 135 EAYAVTAPAPAGWS---AMRLISMQIAFHSYVFGGLTGAVALLELPF-EIKGLEFALC 189

DB 110 ESYGVLLGEALHKKVSPSW-MEGNNVMSYLTWISTIGTLGSTITPNPEMFGDLFALV 168

QY 190 SLFVTLTIDSCRTKQIPSLLAGLSTFIALVTPQALPALLIFIGLTI 241

DB 169 AMFTGLFV-----FOLFGMLSDGKRLVYVIVLASVGLSTFLATPLSGALSV 214

Search completed: April 16, 2004, 06:53:11
 Job time : 73.6156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 23.716 Seconds
(without alignments)

545.110 Million cell updates/sec

Title: US-10-608-504-3

Sequence: 1269 1 MKOTGEIHSSLEVSPEKAL.....LIIFGLLTIRYFLGKAAK 251

Scoring table: BLOSUM62 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep:.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	US-09-471-803A-3	Sequence 3, Appl
2	221.5	17.5	247	US-09-489-039A-8478	Sequence 8778, Ap
3	190	15.0	266	US-09-328-352-6525	Sequence 6525, Ap
4	189.5	14.9	263	US-09-328-352-4486	Sequence 4486, Ap
5	185.5	14.6	255	US-09-107-532A-4920	Sequence 4920, Ap
6	181	14.3	251	US-09-489-039A-10187	Sequence 10187, A
7	175	13.8	345	US-09-252-991A-27947	Sequence 27947, A
8	171.5	13.5	288	US-09-489-039A-12218	Sequence 12218, A
9	161.5	12.7	242	US-09-134-001C-5154	Sequence 5154, Ap
10	141.5	11.2	247	US-09-328-352-6024	Sequence 6024, Ap
11	137	10.8	206	US-09-134-000C-6789	Sequence 6789, Ap
12	134.5	10.6	230	US-09-543-681A-4173	Sequence 4173, Ap
13	113.5	8.9	484	US-09-253-991A-32299	Sequence 32299, A
14	106	8.4	270	US-09-253-991A-18185	Sequence 18185, A
15	104	8.2	555	US-09-253-991A-31521	Sequence 31521, A
16	102.5	8.1	516	US-09-253-991A-21329	Sequence 21329, A
17	100	7.9	506	US-09-540-236-2360	Sequence 2360, Ap
18	98.5	7.8	642	US-08-706-936-2	Sequence 2, Appl
19	98.5	7.8	643	US-08-616-844-39	Sequence 39, Appl
20	98.5	7.8	643	US-08-599-654-39	Sequence 39, Appl
21	98.5	7.8	643	US-08-944-868A-39	Sequence 39, Appl
22	98.5	7.8	643	US-08-944-423A-39	Sequence 39, Appl
23	98.5	7.8	643	US-08-944-486-39	Sequence 39, Appl
24	97	7.6	345	US-09-543-681A-7546	Sequence 7546, Ap
25	96.5	7.6	417	US-08-553-888A-3	Sequence 3, Appl
26	96.5	7.6	490	US-09-253-991A-27210	Sequence 27210, A
27	95.5	7.5	452	US-09-253-991A-19116	Sequence 19116, A

28	95	7.5	1334	2	US-08-996-545-2	Sequence 2, Appl
29	95	7.5	1334	3	US-09-328-320-2	Sequence 2, Appl
30	94	7.4	421	4	US-09-489-039A-7699	Sequence 7699, Ap
31	94	7.4	434	4	US-09-543-681A-7154	Sequence 7154, Ap
32	93.5	7.4	486	4	US-09-291-922-10	Sequence 10, Appl
33	93.5	7.4	591	4	US-09-540-236-2490	Sequence 2490, Ap
34	92.5	7.3	326	4	US-09-107-532A-3665	Sequence 3665, Ap
35	91.5	7.2	274	4	US-09-489-039A-14140	Sequence 14140, A
36	91	7.2	472	4	US-09-489-039A-10315	Sequence 10315, A
37	90.5	7.1	532	4	US-09-252-991A-22421	Sequence 22421, A
38	90	7.1	490	4	US-09-252-991A-30355	Sequence 30355, A
39	88.5	7.0	484	4	US-09-489-039A-11054	Sequence 11054, A
40	88	6.9	2864	4	US-08-469-260A-394	Sequence 394, App
41	88	6.9	2864	4	US-08-467-344A-394	Sequence 394, App
42	88	6.9	2864	4	US-08-467-344A-394	Sequence 394, App
43	87.5	6.9	492	4	US-09-134-001C-4847	Sequence 4847, Ap
44	87.5	6.9	516	4	US-09-489-039A-12893	Sequence 12893, A
45	87.5	6.9	623	4	US-09-252-991A-19867	Sequence 19867, A

ALIGNMENTS

RESULT 1
US-09-471-803A-3
Sequence 3, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMANN

APPLICANT: EGGELE, LOTAR

APPLICANT: PFEFFEL, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS. PROCESS FOR THE

FILE REFERENCE: 21123/265496/MAS

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 3

LENGTH: 251

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

FEATURE:

OTHER INFORMATION: ATCC14752

US-09-471-803A-3

Query Match

Best Local Similarity 100.0%; Score 1269; DB 4; Length 251;

Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKOTGEIHSSLEVSPEKALBPDDGGRYREIAOGLKSLAAGLGMVFIGIAFGILVIOY	60
DB	1	MKOTGEIHSSLEVSPEKALBPDDGGRYREIAOGLKSLAAGLGMVFIGIAFGILVIOY	60
QY	61	GYYWMAAPLFSGLIFAGSITEMLVIALVVGAAFLGAIATLTLLVNRHVFYAFSPFLAYVK	120
DB	61	GYYWMAAPLFSGLIFAGSITEMLVIALVVGAAFLGAIATLTLLVNRHVFYAFSPFLAYVK	120
QY	121	NPIARYSFALIDEXAVTARPPGNGSMRLISQIAFHSYVWGGLGVIAIPIPE	180
DB	121	NPIARYSFALIDEXAVTARPPGNGSMRLISQIAFHSYVWGGLGVIAIPIPE	180
QY	181	IKGLEFALCSLFTVTLTDSCKTKQIPISLLAGLSTFTALVVIPOQAFPAALLIFGLLT	240
DB	181	IKGLEFALCSLFTVTLTDSCKTKQIPISLLAGLSTFTALVVIPOQAFPAALLIFGLLT	240
QY	241	IRYFPLGKAAC 251	
DB	241	IRYFPLGKAAC 251	


```
RESULT 2
US-09-489-039A-8478
; Sequence 8478, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8478
; LENGTH: 247
; TYPE: PRN
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8478

Query Match      17.5%; Score 221.5; DB 4; Length 247;
Best Local Similarity 28.0%; Pred. No. 2.4e-17;
Matches 61; Conservative 41; Mismatches 101; Indels 15; Gaps 6;

QY 32 1AAGLKTSLAAGLGMYPGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAA 91
Db 7 1FRGASAIMPLCIDPFPSPFVIGALVSAGSWQSTAMSAIYAGSACMLALNMUKTGA 66
QY 92 PLGIALTLTLVNRHVFYAFSPFLHYVKNPIARFYS---VFALIDEAVATTAARPAGW 147
Db 67 TLGVATFTTLINRHYVYSASISGTVRE---ASFFKCCFMSYALTDVAVATTKMEGN 123
QY 148 SAMLLI---SMQIAFHSYVWFGULGVAAIAELIP-PEIKGLEFALCSLFTVTLTDSGRTK 203
Db 124 KKEYLFLFGSAMTFWMVWLADFLGALVGSFPIEKYGLDFAMVAFAIIVVPOIKSQ 183
QY 204 KQISLILAGLSTIALVIVP---GQALFALLIFLGL 238
Db 184 ACTVAAVVAASV-GVLLVVLPLYSGLIVASVGLVAGL 220

RESULT 3
US-09-328-352-6525
; Sequence 6525, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6525
; LENGTH: 266
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6525

Query Match      15.0%; Score 190; DB 4; Length 266;
Best Local Similarity 26.6%; Pred. No. 1.1e-13;
Matches 62; Conservative 43; Mismatches 98; Indels 30; Gaps 7;

QY 34 GGLKTSIAGLGMVPIGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAAFL 93
Db 36 QGAKSQALVTLVTVSPFVGSASQFETPEAFLLSCMSWAGASQFLVALLASGSSI 95
QY 94 GAIALTTLLVNRHVFYAFSPFLHYVKNPIARFYSVFAIDENYA---VTAR 143
Db 96 WTLATLVIALDIRHVLVYGPALVNLIPTKLNLKTKAV---WAMGLTDEVFASGMQLSOR 151
```

```
QY 144 PAGSAMLLISMQIAFHSYV---VFQGLTGVAIAELIPPEIKLEPFLCSLFTVTLDS 199
Db 152 RQWSESWMGLISLFSWMSWASGSLGLFPADGVVHLPKFLQALDPLFLFLSLIAA 211
QY 200 CRTKQIPBULLAGLSTIALVIVPQALFPALLIFGLTF-----IRYFFL 246
Db 212 FERKHS---LVAVSLVSVSLACTYINLSAA--IFIGILSGLIAGLFXYYVL 258

RESULT 4
US-09-328-352-4486
; Sequence 4486, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4486
; LENGTH: 263
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4486

Query Match      14.9%; Score 189.5; DB 4; Length 263;
Best Local Similarity 24.0%; Pred. No. 1.3e-13;
Matches 49; Conservative 54; Mismatches 96; Indels 5; Gaps 5;

QY 40 1AAGLGMYPGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAAFLGAILT 99
Db 53 LPLASIVIPMALIASMAVHGLSFKYKALMSGVFAQAQVLSVWMEGASLITTYVT 112
QY 100 TLVNRHVFYAFSPFLHYVKNPIARFYSV-FALIDEAVATTAARPAGMSAMRLISMQIA 158
Db 113 IFPLTAQHFTYALTNRNDISILPLSKRLTGLTDELFAVSPNEKHPOY-LEGAGLC 171
QY 159 FHSYVWFGULGVAAIAELIPPEIK-GLPEFALCSLFTVTLTDSGRTKQIPBULLAGLS-F 216
Db 172 FLFWVSVSLVAILLATLALPMLNTHDPSIATFVAIVMPCKCKPVMAGILMTCVSGF 231
QY 217 TIALVIVPQALFPALL-IFGLL 239
Db 232 VLKFFHIEGALLISGLHGMFIAVI 255

RESULT 5
US-09-107-532A-4920
; Sequence 4920, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
```

CURRENT FILING DATE: 2000-01-27

; CURRENT FILING DATE: 2000-01-27

```

; CURRENT APPLICATION NUMBER: US/03/489,039A
; CURRENT FILING DATE: 2000-01-27

```

PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12218
LENGTH: 288
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12218

Query Match 13.5%; Score 171.5; DB 4; Length 288;
Best Local Similarity 27.5%; Pred. No. 1.8e-11;
Matches 58; Conservative 41; Mismatches 87; Indels 25; Gaps 10;

QY 40 LAAGLMYPIGIAFGILVIOGYEWMAPLFGSLIPAGSTEMLVIALVYGAAPLGAIAL 98
DB 90 LAAGV---VGMVYGLAAVYGFPLWVPLLSVLAASSEFFIGIVASGSPILAA-AA 144
QY 99 TTVLVFRHVFYAFSEPLHVNKNPIAR---FYSVEALIDEAYAVTAAP--GMSAMR 151
DB 145 AGILVVARHIFGVY---VRELVGRRALSFICGHIMDESVFGLSQPTPAQRKAAVW 199
QY 152 LISMQIAFHSYVWPGGLTGAIAELIP-EIKGLEFALCSLFTLLDSCRKQIIPSL 210
DB 200 LCGAGVAL--IWPGLTGAAGVGLPAPETIGLDVFPAILALVPAFKR---TTLI 254
QY 211 LAGLSFTIALVPIGQAL-FAALLIFGLLT 240
DB 255 RAASGALALAAVFPVPTGLPVLISLFGILS 285

RESULT 9

US-09-134-001C-5154
Sequence 5154, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5154
LENGTH: 242
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5154

Query Match 12.7%; Score 161.5; DB 4; Length 242;
Best Local Similarity 24.8%; Pred. No. 2e-10;
Matches 57; Conservative 40; Mismatches 90; Indels 43; Gaps 8;

QY 34 OGLTSLAAGIGMPYIGIAFGILVIOGYEWMAPLFGSLIPAGSTEMLVIALVGAAPL 93
DB 20 QGVEECIPTLLGYAGVGISFGIVAAQNSFVLEITLLCLLIYAAQAQPIITLVIACTPI 79
QY 94 GAILTTLVFNFRHVFYAFSEPLHVNKNPIARFYSV-----ALIDEAYAVTAAP-- 144
DB 80 SAIVTLTIVISR-----MFLSWTLAPNPKQYGFNRVGLTTLTDETPGV-ATIPYV 132
QY 145 -----AGMSAMRLISMQIAFHSYVWPGGLTGAIAELIP-EIKGLEFALCSLFTLLT 197
DB 133 KGEKINDRM-----LHGNIATVAFWTWSCVIGALFGESITSPDALGDFATPAFIFLCI 188
QY 138 DSCRTRKQ-----IPSLLAGLSFTIALVPIGQALFAALL 233
DB 189 SQFEIKTSRLRIYVIVLCVIVMMLLSIIPSVAILI--AAIVALL 236

RESULT 10

US-09-328-352-6024
Sequence 6024, Application US/09328352
Patent No. 6562938
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-039A
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6024
LENGTH: 247
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-6024

Query Match 11.2%; Score 141.5; DB 4; Length 247;
Best Local Similarity 26.4%; Pred. No. 4.1e-08;
Matches 52; Conservative 37; Mismatches 95; Indels 13; Gaps 8;

QY 49 IGIAFGILVIOGYEWMAPLFGSLIPAGSTEMLVIA-LVGAAPLGAIALTTLLVNRH 107
DB 53 VGMISGLAASYGLALMIPCLSLFVLAGTAEFFIFIGFLAVGSPILAA-AIAGLVLNRH 111
QY 108 VFYAFSEPLHVNKNPIARFYSVFPALIDEAYAVTAAP--AGMSAMRLISMQIAFHSYV 164
DB 112 LPFGIAVN-EMIRKFSQFSGHIMDESVLFGVACDPFETKKAAYWLCGIGILS--WP 168
QY 165 FGGLTGAIAELIP-EIKGLEFALCSLFTLLDSCRKQIIPSLLAGLSFTIALVVI 223
DB 169 LGTVAGYIGSALIPDKTFGDLAIFPAILLITSALKR---VTRKAAFRGSLALITT 225
QY 224 PGQALFAALLIFL-GLT 239
DB 226 PELASGLPILSLFLGI 242

RESULT 11

US-09-134-000C-6789
Sequence 6789, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SEQ ID NO 6789
SOFTWARE: Patentin version 3.1
SEQ ID NO 6789
LENGTH: 206
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-6789

Query Match 10.8%; Score 137; DB 4; Length 206;
Best Local Similarity 26.8%; Pred. No. 1e-07; Indels 24; Gaps 9;
Matches 52; Conservative 42; Mismatches 76;

QY 71 SGLIPAGSTEMLVIALVGAAPLGAIALTTLLVNRHVFYAFS-PLHVNKNPIARFYSV 129
DB 17 SFPIFAGSAQFVTVSMILGSGPILSLIVLATFIVNARMIMGMTIAPYFKAESLGRMLWG 76
QY 130 PALIDEAYAV-----TAAP--GMSAMRLISMQIAFHSYVWPGGLTGAIAELIP- 179
DB 77 TLLTDESFLGKMKLNHTGRUSFEWFNANLIS-----YATVWSTTIGAVLGRFIANP 131
QY 180 EIKGLEFALCSLFLV---TLLDSCRKQIIPSLLAGLSF---TIALVPIGQAL-FAAL 232

US-09-252-991A-31521
; Sequence 31521, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31521
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31521

Query Match 8.2%; Score 104; DB 4; Length 555;
Best Local Similarity 22.1%; Pred. No. 0.0028;

Matches 75; Conservative 32; Mismatches 92; Indels 140; Gaps 16;

QY 32 IAQGLKTSIAAGLGMPI---GIAFGILVYQYGYEWMAAPLFSGL-----IFAGSTEML 82
DB 216 ICMGLVALLLPWIGRWLAERHGIYCGILAIAAAC-LMRELDGGLALIASALAGSGVAI 274
QY 83 VIALVGA-----APLGAIALTT 100
DB 275 IOMALPEGVVGRMPFRVPAMGLYSASIMAGCGTAIVLSPRIAEHFSNMQAGANAVPA 334
QY 101 LVNFRHVFYAFSPPLHV---KNPIARFYS-----VPALIDEAYAVTA----- 142
DB 335 LLA---LLMWFAPRPREVLPSAGEPVPHFPGNRGWLAVYFGLINGYTSWVAMLPHY 391
QY 143 -RPAQMSA-----WRLISMOIAFHSYWFVGG 168
DB 392 YRQLSMSHDSGGVIGMTTFQVLAALSVPLIRRLDRRPRMLAALVQ-----LGGF 445
QY 169 TGVAAIPEIPEIKGLEFAL-----C-SLEFVTLTDSCKTKQIPSL--LLAGLSFTI 218
DB 446 CGIL---LMPQHAALWVALIGYGLGACFALSTLTLDHLHEPRPAGSLAAFGIGIFII 502
QY 219 ALVY-----IPGQ-----ALFALLIFGLITRY 243
DB 503 TGIVPYLTGMLRDVYTSFQASWMLLAASVAMLVTLRF 541

Search completed: April 16, 2004, 06:56:34
Job time : 24.7716 secs

Tue Apr 20 06:47:18 2004

us-10-608-504-3.rapn

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:53:52 ; Search time 17.4791 Seconds
(without alignments)
386.360 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269
Sequence: 1 MCKQEHHSLEVSRSKAL.....LIFLGLITVFLGKAK 251

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193692 seqs, 26905285 residues

Total number of hits satisfying chosen parameters: 193692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185.5	14.6	255	US-10-417-884A-4920	Sequence 4920, Ap
2	161.5	12.7	242	US-10-724-972A-6721	Sequence 6721, Ap
3	99	7.8	455	US-60-556-841-2480	Sequence 2480, Ap
4	98	7.7	477	US-60-556-841-7285	Sequence 7285, Ap
5	98	7.7	760	PCT-US04-03417-32	Sequence 32, Appl
6	98	7.7	760	US-10-772-636-32	Sequence 32, Appl
7	98	7.7	858	US-10-725-081-14	Sequence 14, Appl
8	98	7.7	858	US-60-554-751-14	Sequence 14, Appl
9	96.5	7.6	417	US-09-857-097C-148	Sequence 148, Appl
10	96.5	7.6	417	US-09-857-097C-149	Sequence 149, Appl
11	96.5	7.6	417	US-09-857-097C-150	Sequence 150, Appl
12	96.5	7.6	417	US-09-857-097C-152	Sequence 152, Appl
13	92.5	7.3	336	US-10-417-884A-3665	Sequence 3665, Ap
14	91.5	7.2	389	US-60-556-841-2463	Sequence 2463, Ap
15	90	7.1	417	US-09-857-097C-151	Sequence 151, Appl
16	90	7.1	560	US-60-556-841-2332	Sequence 2332, Ap
17	89.5	7.1	552	US-10-781-014-654	Sequence 654, Appl
18	88.5	7.0	144	US-10-767-701-61776	Sequence 61776, A
19	88	6.9	471	US-60-556-841-4957	Sequence 4957, Ap
20	88	6.9	476	US-60-556-841-12104	Sequence 12104, A
21	87.5	6.9	492	US-10-724-972A-6373	Sequence 6373, Ap
22	87	6.9	497	US-60-556-841-10192	Sequence 10192, A
23	87	6.9	508	US-60-556-841-6098	Sequence 6098, Ap
24	87	6.9	555	US-10-100-683-7076	Sequence 7076, Ap
25	86.5	6.8	306	US-10-275-027A-322	Sequence 322, Appl
26	86	6.8	850	US-10-725-081-4	Sequence 4, Appl

27	86	6.8	852	US-60-554-751-12	Sequence 12, Appl
28	85.5	6.7	235	US-10-417-884A-3775	Sequence 3775, Ap
29	85.5	6.7	476	US-60-556-841-8474	Sequence 8474, Ap
30	85	6.7	392	US-60-556-841-4013	Sequence 4013, Ap
31	84.5	6.7	388	PCT-US04-07467-564	Sequence 564, Appl
32	84.5	6.7	426	US-60-556-841-6894	Sequence 6894, Ap
33	84	6.6	471	US-60-556-841-3188	Sequence 3188, Ap
34	84	6.6	858	US-60-554-751-13	Sequence 13, Appl
35	83.5	6.6	382	US-60-556-841-403	Sequence 403, Appl
36	83.5	6.6	453	US-60-556-841-6300	Sequence 6300, Ap
37	83	6.5	465	US-60-556-841-8829	Sequence 8829, Ap
38	82.5	6.5	305	US-10-417-884A-4215	Sequence 4215, Ap
39	82.5	6.5	375	US-10-767-701-43678	Sequence 43678, A
40	82	6.5	304	US-60-556-841-1622	Sequence 1622, Ap
41	82	6.5	454	US-60-556-841-2905	Sequence 2905, Ap
42	82	6.5	464	US-60-556-841-8095	Sequence 8095, Ap
43	82	6.5	469	US-60-556-841-11213	Sequence 11213, A
44	82	6.5	634	US-09-830-915D-37	Sequence 37, Appl
45	82	6.5	675	US-60-551-121-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-417-884A-4920
; Sequence 4920, Application US/10417884A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/417, 884A
; FILING DATE: 17-Apr-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107, 532
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085, 598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40, 489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4920:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...255

SEQUENCE DESCRIPTION: SEQ ID NO: 4920:
US-10-417-884A-4920

Query Match 14.6%; Score 185.5; DB 6; Length 255;
Best Local Similarity 25.0%; Pred. No. 8.1e-12;
Matches 63; Conservative 51; Mismatches 105; Indels 33; Gaps 10;

16 SKALEDDEKGYRREYIAQGLKSLAAGMYPGIAFGILVYQYEWMAAPLPSGLIF 75
15 AKAVLLIDK-----LDKTKAIKOTLPFTVFGYIGLAFGVGAAGHPPLVTTLSLVY 70
76 AGSTEMLVIALVGAAPLGAIALTTLVNERHVFYAFSPPLHVKNPPIARFYSVPALI-D 134
71 AGSAQFTVWMLASHPSPLSLVSTPLVNSRMILMSWTIAPYFKKRLQLNLLIGTLTD 130
135 EAYAVTPARAGWSAKRLISMOIAFH-----SY-WYFGGLTGVAAIAELIPEIK-G 183
131 ESPAL-----GNNKLNVTGQKLNFRWMTANNISYLTWVGSSLVGALLGNFTTDPKKEG 184
184 LEFALCSLFTL-----TLDSCKRTKQIPSLLAGLFT--IALVYIPOGALFALLIF- 235
185 LDRIVAMFTGLVLYQYISDRNTSKRL-QLILIGTLILVYGLIFIPSNLVIVVVTLLG 243
236 --IGLITIRYF 245
244 CGLGVWIKHAF 255

RESULT 2
US-10-724-972A-6721

Sequence 6721, Application US/10724972A
GENERAL INFORMATION:
APPLICANT: Doucette-Stamm, Lynn
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: PATHO3-16
CURRENT APPLICATION NUMBER: US/10/724, 972A
CURRENT FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: 09/450, 969
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: 09/134, 001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 7544
SEQ ID NO 6721
LENGTH: 242
TYPE: PRT
ORGANISM: S. epidermidis
US-10-724-972A-6721

Query Match 12.7%; Score 161.5; DB 6; Length 242;

Best Local Similarity 24.8%; Pred. No. 2.4e-09;
Matches 57; Conservative 40; Mismatches 90; Indels 43; Gaps 8;

34 GGLKTSIALGLMYPGIAFGILVYQYEWMAAPLPSGLIFAGSTEMLVIALVGAAPL 93
20 QGVCECPITLLGAVAGLSFGIYASQNSFVLEITLCLITVYGAQFICTIVIGTPI 79
94 GATALLTLVFRHVFYAFSPPLHVKNPPIARFYSVP-----ALIDEAVATAARP-- 144
80 SAIVLTLLVNSR-----WFLSMTLAPYKQGFENRVLGLTLLTDEFQV-ALTTPV 132
145 -----AGNSAKRLISMOIAFHSHYVWFGILTVAAIALIFP-EIKGLEFALCSLFTLLD 197
133 KGEKINDR-----LHGNTAVYLFWTVCVIGALFGYISNPALGIDFAITGMFTFLCI 188
198 DSCRTKQ-----IPSLLAGLFTIALVYIPOGALFALL 233
189 SQFGIKSKRLRIYIVIVCVIWMLLSLPEYVAIIL--AIVAAAL 236

RESULT 3

US-60-556-841-2490
Sequence 2490, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556, 841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 2490
LENGTH: 455
TYPE: PRT
ORGANISM: Bacillus halodurans
US-60-556-841-2490

Query Match 7.8%; Score 99; DB 7; Length 455;
Best Local Similarity 27.5%; Pred. No. 0.017;
Matches 56; Conservative 26; Mismatches 68; Indels 54; Gaps 11;

45 GMYPGIAFGILVYQYEWMAAPLPSGLIFAGSTEMLVIALVY-----GAAPLGAIALTT 100
191 GLFPEGF-LGL-----WPAIFGYAFAGGIEIGLILANRLEEKGAAGKGRAML-- 238
101 LVNERHVFYAFSPPLHVKNPPIARFYSVPALIDEAVATTAAPAGWSAKRLISMOIAFH 160
239 LLTF-----LVVVSGLAILVPLNRFT-----ETSPFITAKRYG-----LPIVSH 282
161 SYWVFGILTV-----IAELIPEIKGLEFALCSLFTLLDSCKRTKQIPSLLAGLS 215
283 --LNGVLILAGSTWASL-----FAVTSILVTLAAD-----KQAPNLFARKVK 325
216 FTIALVYIPOGALFALLIFGL 239
326 NKIALPALGLRIAGLALSVLSTL 349

RESULT 4

US-60-556-841-7285
Sequence 7285, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556, 841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 7285
LENGTH: 477
TYPE: PRT
ORGANISM: Ferrioplasma acidarmanus
US-60-556-841-7285

Query Match 7.7%; Score 98; DB 7; Length 477;

Best Local Similarity 24.6%; Pred. No. 0.023;
Matches 55; Conservative 31; Mismatches 86; Indels 52; Gaps 10;

53 FGLVYQYEW-----AAPLPSGLIFAGSTEMLVIALVGA---APLGAIALTTL 101
92 YGLII-----GYMWSYFVVTNALIFTSQVAPIPGTISYFLHFNQSYTWIPI---MTAF 144
102 LVNERHVFY-----AFSPPLHVKNPPIARFYSVPALI-----DEAVATAAPAGWSAW 150
145 IYIVLVIAVLGKPSLASFASVIERALVITSVIIVVAGSKNTVPTFAGADGPAA 204
151 RLISMOIAFHSHYVWFGILTVAAIAE-----LIFEIKGLEFALCSLFTV-- 194
205 VAVGVMTAIFS--MSGSAVVTLGEEAQPRKNIKALLSFIITGVFVLTSTALTVGW 262
195 -LTDSCKRTKQIPSLLAGL-----G-SFTIALVYIPOGALFALL 233

Tue Apr 20 06:47:18 2004

us-10-608-504-3.rapn

Page 3

Db 263 GVSXMSFTTSSVPLIVSDKYLGPFTIVLFVFIINSIFAGSL 306

RESULT 5
PCT-US04-03417-32

Sequence 32, Application PC/TUS0403417

GENERAL INFORMATION:

APPLICANT: Kelly, Louise M.

APPLICANT: Carroll, Joseph M.

APPLICANT: Farlow, Deborah

APPLICANT: Healy, Aileen

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,

TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,

TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,

TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,

TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR

TITLE OF INVENTION: 5014

FILE REFERENCE: MP103-015

CURRENT APPLICATION NUMBER: PCT/US04/03417

PRIOR FILING DATE: 2004-02-05

PRIOR APPLICATION NUMBER: US 60/445,241

PRIOR FILING DATE: 2003-02-05

PRIOR APPLICATION NUMBER: US 60/448,389

PRIOR FILING DATE: 2003-02-18

PRIOR APPLICATION NUMBER: US 60/456,320

PRIOR FILING DATE: 2003-03-20

PRIOR APPLICATION NUMBER: US 60/460,279

PRIOR FILING DATE: 2003-04-03

PRIOR APPLICATION NUMBER: US 60/465,924

PRIOR FILING DATE: 2003-04-28

PRIOR APPLICATION NUMBER: US 60/470,052

PRIOR FILING DATE: 2003-05-13

PRIOR APPLICATION NUMBER: US 60/498,106

PRIOR FILING DATE: 2003-08-26

PRIOR APPLICATION NUMBER: US 60/500,179

PRIOR FILING DATE: 2003-09-04

PRIOR APPLICATION NUMBER: US 60/502,909

PRIOR FILING DATE: 2003-09-15

PRIOR APPLICATION NUMBER: US 60/510,351

PRIOR FILING DATE: 2003-10-10

Remaining Prior Application data removed - See file wrapper or PALM.

NUMBER OF SEQ ID NOS: 80

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 32

LENGTH: 760

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US04-03417-32

Query Match 7.7%; Score 98; DB 1; Length 760;

Best Local Similarity 25.6%; Pred. No. 0.042;

Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

Db 80 EMVLIVLVG-----AALGALATLLVNFHVFYAFSPPLHVNKNPIARFVSVPALD 134

Db 254 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--YFPLKTL-----WRSFFALV 300

Qy 135 EAYAVTAAAPAGSAMRLISMOIAFHSYVWFGGLTGVAAELIPEIKGLEFALC-SLFV 193

Db 301 AAFLLRSINPFGNS-RLVLFYVEYHTPMY-----MAELPFIILGVFGMLGTLFI 350

Qy 194 TLTLDSGRTK-----QIPSLLAGLSFTTALVIP 224

Db 351 RCNIAMCRRKRTKRLGKXPVLEIVVATATIAVP 386

Db 351 RCNIAMCRRKRTKRLGKXPVLEIVVATATIAVP 386

RESULT 6

US-10-772-636-32

Sequence 32, Application US/10772636

GENERAL INFORMATION:

APPLICANT: Kelly, Louise M.

APPLICANT: Carroll, Joseph M.

APPLICANT: Farlow, Deborah

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,

TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,

TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,

TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,

TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR

TITLE OF INVENTION: 5014

FILE REFERENCE: MP103-015P1RKNXNIM

CURRENT APPLICATION NUMBER: US/10/772,636

PRIOR FILING DATE: 2004-02-05

PRIOR APPLICATION NUMBER: US 60/445,241

PRIOR FILING DATE: 2003-02-05

PRIOR APPLICATION NUMBER: US 60/448,389

PRIOR FILING DATE: 2003-02-18

PRIOR APPLICATION NUMBER: US 60/456,320

PRIOR FILING DATE: 2003-03-20

PRIOR APPLICATION NUMBER: US 60/460,279

PRIOR FILING DATE: 2003-04-03

PRIOR APPLICATION NUMBER: US 60/465,924

PRIOR FILING DATE: 2003-04-28

PRIOR APPLICATION NUMBER: US 60/470,052

PRIOR FILING DATE: 2003-05-13

PRIOR APPLICATION NUMBER: US 60/498,106

PRIOR FILING DATE: 2003-08-26

PRIOR APPLICATION NUMBER: US 60/500,179

PRIOR FILING DATE: 2003-09-04

PRIOR APPLICATION NUMBER: US 60/502,909

PRIOR FILING DATE: 2003-09-15

PRIOR APPLICATION NUMBER: US 60/510,351

PRIOR FILING DATE: 2003-10-10

Remaining Prior Application data removed - See file wrapper or PALM.

NUMBER OF SEQ ID NOS: 80

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 32

LENGTH: 760

TYPE: PRT

ORGANISM: Homo sapiens

US-10-772-636-32

Query Match 7.7%; Score 98; DB 6; Length 760;

Best Local Similarity 25.6%; Pred. No. 0.042;

Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

Db 80 EMVLIVLVG-----AALGALATLLVNFHVFYAFSPPLHVNKNPIARFVSVPALD 134

Db 254 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--YFPLKTL-----WRSFFALV 300

Qy 135 EAYAVTAAAPAGSAMRLISMOIAFHSYVWFGGLTGVAAELIPEIKGLEFALC-SLFV 193

Db 301 AAFLLRSINPFGNS-RLVLFYVEYHTPMY-----MAELPFIILGVFGMLGTLFI 350

Qy 194 TLTLDSGRTK-----QIPSLLAGLSFTTALVIP 224

Db 351 RCNIAMCRRKRTKRLGKXPVLEIVVATATIAVP 386

Db 351 RCNIAMCRRKRTKRLGKXPVLEIVVATATIAVP 386

RESULT 7

US-10-725-081-14

Sequence 14, Application US/10725081

GENERAL INFORMATION:

APPLICANT: ADLER, JON ELLIOT

APPLICANT: ZOSULYA, SERGEY

APPLICANT: LI, XINDONG SHAN

APPLICANT: O'CONNELL, SHAN

APPLICANT: STASZEWSKI, LENA

TITLE OF INVENTION: TLR TASTE RECEPTORS AND GENES ENCODING SAME

FILE REFERENCE: 078003/0277870/RXT

CURRENT APPLICATION NUMBER: US/10/725,081

PRIOR FILING DATE: 2003-12-02

PRIOR APPLICATION NUMBER: 60/187,546

PRIOR FILING DATE: 2000-03-07

PRIOR APPLICATION NUMBER: 60/195,536
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: 60/209,840
 PRIOR FILING DATE: 2000-06-06
 PRIOR APPLICATION NUMBER: 60/214,213
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 30/226,448
 PRIOR FILING DATE: 2000-08-17
 PRIOR APPLICATION NUMBER: 60/259,227
 PRIOR FILING DATE: 2001-01-03
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 14
 LENGTH: 858
 TYPE: PRT
 ORGANISM: Rattus sp.
 US-10-725-081-14

Query Match 7.7%; Score 98; DB 6; Length 858;
 Best Local Similarity 22.3%; Pred. No. 0.048;
 Matches 59; Conservative 35; Mismatches 92; Indels 78; Gaps 12;

QY 52 AREGLVIOGYEWMAAPLPS-----GLIFAGSTEMLV-----ALVGAAPLG 94
 DB 591 ALGLFV-----HYWDSPLVQASGSLFCFGLICGLFCLSVLFPGRPSASCLAQQPM 645
 QY 95 AIALT-----TLVNFPRHVFYAFSPFL-----HVKNPFIARFVSFALIDE---AVAV 139
 DB 646 HLPLTGCLSTFLQAAEIFESELPLSMANWLCSYRGWAMLVLLATLVEALCAWYL 705
 QY 140 TAARPAGSAMRLISMQIAFH---SYWVEGGLTV---ATAELIFP-----E 180
 DB 706 MAPPEVVTMQVLPTEVLEHCHMRSMWSLGLVHINNAVLAFICFGLTFLVQSQRKYNR 765
 QY 181 IKGLFALCSLFTVTLTDSCKTKQIPSLLAGLSTFIALVIFGALPAAL----- 232
 DB 766 ARGTFAMLAYFIWV-----SFVP---LLANVQVAPVQVGMALIFCALGIATFHL 816
 QY 233 -----LIFGLTLTRFPLGKAK 251
 DB 817 PKCYVLMPELNTQEFFLGSRPK 840

RESULT 8
 US-60-554-751-14
 Sequence 14, Application US/60554751
 GENERAL INFORMATION:
 APPLICANT: L1, Xia
 APPLICANT: L1, Weihua
 APPLICANT: Reed, Danielle R.
 APPLICANT: Bachmanov, Alexander A.
 APPLICANT: Brand, Joseph G.
 TITLE OF INVENTION: TASTE RECEPTOR OF THE T1R FAMILY FROM DOMESTIC CAT
 FILE REFERENCE: NON-0298
 CURRENT APPLICATION NUMBER: US/60/554,751
 CURRENT FILING DATE: 2004-03-19
 NUMBER OF SEQ ID NOS: 98
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 14
 LENGTH: 858
 TYPE: PRT
 ORGANISM: Rattus rattus
 US-60-554-751-14

Query Match 7.7%; Score 98; DB 7; Length 858;
 Best Local Similarity 22.3%; Pred. No. 0.048;
 Matches 59; Conservative 35; Mismatches 92; Indels 78; Gaps 12;
 QY 52 AREGLVIOGYEWMAAPLPS-----GLIFAGSTEMLV-----ALVGAAPLG 94
 DB 591 ALGLFV-----HYWDSPLVQASGSLFCFGLICGLFCLSVLFPGRPSASCLAQQPM 645
 QY 95 AIALT-----TLVNFPRHVFYAFSPFL-----HVKNPFIARFVSFALIDE---AVAV 139

DB 646 HLPLTGCLSTFLQAAEIFESELPLSMANWLCSYRGWAMLVLLATLVEALCAWYL 705
 QY 140 TAARPAGSAMRLISMQIAFH---SYWVEGGLTV---ATAELIFP-----E 180
 DB 706 MAPPEVVTMQVLPTEVLEHCHMRSMWSLGLVHINNAVLAFICFGLTFLVQSQRKYNR 765
 QY 181 IKGLFALCSLFTVTLTDSCKTKQIPSLLAGLSTFIALVIFGALPAAL----- 232
 DB 766 ARGTFAMLAYFIWV-----SFVP---LLANVQVAPVQVGMALIFCALGIATFHL 816
 QY 233 -----LIFGLTLTRFPLGKAK 251
 DB 817 PKCYVLMPELNTQEFFLGSRPK 840

RESULT 9
 US-09-857-097C-148
 Sequence 148, Application US/09857097C
 GENERAL INFORMATION:
 APPLICANT: Urbanek, Stanislaw J.
 APPLICANT: Barker, Robert, N.
 TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPITOPES
 FILE REFERENCE: P097
 CURRENT APPLICATION NUMBER: US/09/857,097C
 CURRENT FILING DATE: 2001-05-31
 PRIOR APPLICATION NUMBER: 9826378.3
 PRIOR FILING DATE: 1998-12-01
 NUMBER OF SEQ ID NOS: 152
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 148
 LENGTH: 417
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Rhce Residues 111-125
 US-09-857-097C-148

Query Match 7.6%; Score 96.5; DB 5; Length 417;
 Best Local Similarity 21.7%; Pred. No. 0.028;
 Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;
 QY 18 AALEPDDKG-YRREYIAQGLKTSIAGLGMP-----IGIAFGLVIOGYEWMAAP 68
 DB 36 ASLE-DQKGLVASYQVGDITVMAAIGLFTSSFRHSWSVAENLMTALGVQW--AI 92
 QY 69 LPSC-----LIPAGST-----EMLVIALVWGAAPLGA 95
 DB 93 LIDGFLSPGSKVITLFSRLATMGAMSVLISAGAVLGKVNLAQVWVIVETALGT 152
 QY 96 IAL-----TLVNFPRHVFYAFS--FPLHV---VKNPI-----ARFVSVALI 133
 DB 133 LRMVSNIFNTDYHNNLRH--FYVFAAYFGLVAVWCLEPKGTEDNDORATIPSLAML 211
 QY 134 DEAYAVTARAPAGSAMRLISMQI-----AFSYWVFGGLTVGAIAELIPEIKG 183
 DB 212 GALEL-----WMPFVSNSPLRSPLOKKNAMNTY-----ALAVSVYTKISG 255
 QY 184 LEFALCSLFTVTLTDSCKTKQIPSLILA--GLSEFTIALVIFGALPAALIFGLIT 241
 DB 256 SLEAPORRKSMT-----YHSAVLAVGAVAGTSCHLIPSWLMLVGLVAGLISI 306

RESULT 10
 US-09-857-097C-149
 Sequence 149, Application US/09857097C
 GENERAL INFORMATION:
 APPLICANT: Urbanek, Stanislaw J.
 APPLICANT: Barker, Robert, N.
 TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPITOPES
 FILE REFERENCE: P097
 CURRENT APPLICATION NUMBER: US/09/857,097C
 CURRENT FILING DATE: 2001-05-31

PRIOR APPLICATION NUMBER: 9826378.3
PRIOR FILING DATE: 1998-12-01
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 149
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 121-135
US-09-857-097C-149

Query Match
Best Local Similarity 21.7%; Pred. No. 0.028; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 AALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVGDLTVMALIGLGLTSSFRHSSVAFVFLFALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVGAAPLGA 95
DB 93 LLDGFLSQPBGKVVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVVAVLEVTALGT 152
QY 96 IAL-----TTLVNFHVFYAFS--FPLHV--VKNPI-----ARFVSVALI 133
DB 153 LRMVINSIFNTDYHNLNRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGWSAMRLISMOI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRSPIQRKNAMFNTRY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLILA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPORKISMT-----YHSAVLAGVAVGTSCHLIPSPWLAMVLGVAGLISI 306

RESULT 11
US-09-857-097C-150
Sequence 150; Application US/09857097C
GENERAL INFORMATION:
APPLICANT: Urbanik, Stanislaw J.
APPLICANT: Barker, Robert, N.
TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPTTOPES
FILE REFERENCE: P097
CURRENT APPLICATION NUMBER: US/09/857,097C
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 9826378.3
PRIOR FILING DATE: 1998-12-01
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 150
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 131-145
US-09-857-097C-150

Query Match
Best Local Similarity 7.6%; Score 96.5; DB 5; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 AALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVGDLTVMALIGLGLTSSFRHSSVAFVFLFALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVGAAPLGA 95
DB 93 LLDGFLSQPBGKVVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVVAVLEVTALGT 152
QY 96 IAL-----TTLVNFHVFYAFS--FPLHV--VKNPI-----ARFVSVALI 133

DB 153 LRMVINSIFNTDYHNLNRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGWSAMRLISMOI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRSPIQRKNAMFNTRY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLILA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPORKISMT-----YHSAVLAGVAVGTSCHLIPSPWLAMVLGVAGLISI 306

RESULT 12
US-09-857-097C-152
Sequence 152; Application US/09857097C
GENERAL INFORMATION:
APPLICANT: Urbanik, Stanislaw J.
APPLICANT: Barker, Robert, N.
TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPTTOPES
FILE REFERENCE: P097
CURRENT APPLICATION NUMBER: US/09/857,097C
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 9826378.3
PRIOR FILING DATE: 1998-12-01
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 152
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 151-165
US-09-857-097C-152

Query Match
Best Local Similarity 21.7%; Pred. No. 0.028; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 AALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVGDLTVMALIGLGLTSSFRHSSVAFVFLFALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVGAAPLGA 95
DB 93 LLDGFLSQPBGKVVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVVAVLEVTALGT 152
QY 96 IAL-----TTLVNFHVFYAFS--FPLHV--VKNPI-----ARFVSVALI 133
DB 153 LRMVINSIFNTDYHNLNRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGWSAMRLISMOI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRSPIQRKNAMFNTRY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLILA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPORKISMT-----YHSAVLAGVAVGTSCHLIPSPWLAMVLGVAGLISI 306

RESULT 13
US-10-417-884A-3665
Sequence 3665; Application US/10417884A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/417, 884A

FILING DATE: 17-Apr-2003

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/107, 532

FILING DATE: 30-Jun-1998

APPLICATION NUMBER: 60/085, 598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arieliello, Pamela Deneke

REGISTRATION NUMBER: 40, 489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3665:

SEQUENCE CHARACTERISTICS:

LENGTH: 326 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...326

SEQUENCE DESCRIPTION: SEQ ID NO: 3665:

US-10-417-884A-3665

Query Match

Best Local Similarity 24.6%; Score 92.5; DB 6; Length 326;

Matches 34; Conservative 24; Mismatches 45; Indels 35; Gaps 6;

Db

143 RPAGWSAW-RLSMQIAHSYVFGLTGVAIAIIPFETKLEPAL-----CSLFV 193

6 RVDWRYGRVAVLALVILSISLFGVSNVQLSRLDQSGSLILVTRLPRTICILV 65

134 TLTLDSC-----RRKQIPSLLAG-----LSFTIALVIRPGAL-----FA 230

66 GATSSICGIMQHILQNFVSPTTAGTMDSARLGLIVAMIFLPQASLIVRSLTAFCAFA 125

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Db

231 ALLIFGLITIRYFPLGK 248

126 GTLLFSLTRL--FLQK 140

Query Match

Best Local Similarity 21.5%; Score 90; DB 5; Length 417;

Matches 69; Conservative 36; Mismatches 104; Indels 112; Gaps 13;

Db

18 AALEPDKG-YRREIAQGLKTSLAGLAMP-----ICIAFGLVITYGYEWMAAP 68

36 ASLE-DQKGLVASVYQVQDQTLVMAAIGLFTSSFRHSWSVAFTLALGVQW--AI 92

69 LFSGLIFAGSTENLV-----ALVGAAPLGAIALTLV-----N 104

93 LLDGFLSQPSPSKVITLPSIRLATMSALSVLISDAVIGKNLAQVVMVVEVTALCN 152

105 FR-----HVFYAF----- 112

153 LRNVISNIFNTDVMHMMCHIVFAAYFGSVAMCLEPKPLPEGTEEDNDQTATIPISAMLG 212

113 -----SPLHVKVNIARFYSVPALIDEAYNTARPPGWSAMRLSNQ-----IA 158

213 ALFLMTFWSFNALRSPERKNVFE---NTYAAVSVVTAISSSLAHQOKISKTY 269

159 FHSYVWFGGLTGAIAELIPFB-----IKGLEFALCSLFTLTDSCTR-KOIPSLLAG 213

270 VHSANVLAGVAVTSCHLIPSPWLAVLGLVAGLVGAKVLPCCCNVLTGIPHSING 329

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

214 LSFTIALVIRPGALPALII 234

330 YNF--SLDGLGEIIVIVLV 348

Db

Db 555 GGATGATACCAAGAGTTTTCATTTGCTTGGAAATCCCAATTGAAAGTTTTCAG 614
 Qy 194 ATCTGAAGACTTAACTTTGGTCTCATTCGCTTGCATTAAGTGTGGCCATCTTC 253
 Db 615 CTCAGAGCAGATATATGTCATGCTGTATACCGGGGTTCATTTTTCGACTT 674
 Qy 254 TTGGCGGTGACGACCTTGTGAGCGCTTGGCGCTGG 290
 Db 675 TTTTACTTGCTTTGATGCTGAAGCTAGCTGGGG 711

RESULT 15

CK295424

LOCUS

DEFINITION

CK295424 936 bp mRNA linear EST 15-DEC-2003
 EST758138 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NBMCU77 5'
 end, mRNA sequence.

ACCESSION CK295424 GI:39879801
 VERSION CK295424.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

1. 936
 location/Qualifiers
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMCU77"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="MDH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 936;
 Best Local Similarity 47.5%; Fred. No. 26;

Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 74 TTCCGTTCTTAAGCCCTACGTAATCAATTTGGGCAAAATGGCGATG 133
 Db 419 TTCTATGTATATGTCAGCCGAAACATCTCAAGTGGCGGGAATTCGGCTGT 478
 Qy 134 GGATGCCAGCAGAAATCCTTGGCATTTGACCGCATCAAGTTTGGCAGCATGCGATAG 193
 Db 479 GGATGATACCAAGTGTTCATTTGCTTTGGAAATTCCAATTATGAAAGTTTGGCAG 538

Qy 194 ATCTGAAGACTTAACTTTGGTCTCATTCGCTTGCATTAAGTGTGGCCATCTTC 253
 Db 539 CTCAGAGCAGATATATGTCATGCTGTATACCGGGGTTCATTTTTCGACTT 598
 Qy 254 TTGGCGGTGACGACCTTGTGAGCGCTTGGCGCTGG 290
 Db 599 TTTTACTTGCTTTGATGCTGAAGCTAGCTGGGG 635

Search completed: April 17, 2004, 20:01:49
 Job time : 1050.18 secs

QY 74 TTCGGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 404 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGGCGGGGAAATTCGGCTGT 463
QY 134 GGATGCGACAGAGAAATCCTTGCCATTTTGACCGCATCAAGTTTGCGACATGGCATAG 193
DB 464 GGATGATACCAAGAGTGTTCATTGCTTGGAATTCCTCAATTAAGAGTTTGGCAAG 523
QY 194 ATCTGAAGACTCTAACCTTTGGTCTCATTTGCGCTTGCGATTACAGTGGTGGCATCTTC 253
DB 524 CTCAGAGCAAGATTAAGTCATGCTGTATAGCCGGGCTTCATGTTTTCGACTT 583
QY 254 TTGGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGG 290
DB 584 TTTTACTTGCTTTGATGCTGACACTAGGCTGGGG 620

RESULT 13
LOCUS CK288648
DEFINITION 890 bp mRNA linear EST 15-DEC-2003
normalised, full-length Nicotiana benthamiana cDNA clone NEMB102 5' end, mRNA sequence.

ACCESSION CK288648
VERSION CK288648
KEYWORDS CK288648.1 GI:39866370
SOURCE EST.
ORGANISM Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanales; Nicotiana.
1 (bases 1 to 890)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanicheva, S.A., Day, B., Staskawicz, B., Jin, H. and Baker, B.
REFERENCE Generation of EST sequences from Nicotiana benthamiana
TITLE Unpublished (2003)
JOURNAL Contact: Robin Buell
COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES
source
1..890
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMB102"
/tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN
Query Match 10.7%; Score 34.6; DB 14; Length 890;
Best Local Similarity 47.5%; Pred. No. 25;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 543 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGGCGGGGAAATTCGGCTGT 602

QY 134 GGATGCGACAGAGAAATCTTGCATTTTGACCGCATCAAGTTTGCGACATGGCATAG 193
DB 603 GGATGATACCAAGAGTGTTCATTGCTTGGAATTCCTCAATTAAGAGTTTGGCAAG 662
QY 194 ATCTGAAGACTCTAACCTTTGGTCTCATTTGCGCTTGCGATTACAGTGGTGGCATCTTC 253
DB 663 CTCAGAGCAAGATTAAGTCATGCTGTATAGCCGGGCTTCATGTTTTCGACTT 722
QY 254 TTGGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGG 290
DB 723 TTTTACTTGCTTTGATGCTGACACTAGGCTGGGG 759

RESULT 14
LOCUS CK295854
DEFINITION 922 bp mRNA linear EST 15-DEC-2003
normalised, full-length Nicotiana benthamiana cDNA clone NEMC72 5' end, mRNA sequence.

ACCESSION CK295854
VERSION CK295854
KEYWORDS CK295854.1 GI:39880657
SOURCE EST.
ORGANISM Nicotiana benthamiana
Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanales; Nicotiana.
1 (bases 1 to 922)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanicheva, S.A., Day, B., Staskawicz, B., Jin, H. and Baker, B.
REFERENCE Generation of EST sequences from Nicotiana benthamiana
TITLE Unpublished (2003)
JOURNAL Other ESTs: EST758569
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES
source
1..922
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMC72"
/tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN
Query Match 10.7%; Score 34.6; DB 14; Length 922;
Best Local Similarity 47.5%; Pred. No. 26;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 495 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGGCGGGGAAATTCGGCTGT 554
QY 134 GGATGCGACAGAGAAATCTTGCATTTTGACCGCATCAAGTTTTCGACATGGCATAG 193

ORIGIN

/clone="CS0D1052YB18"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was placed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 10.7%; Score 34.6; DB 13; Length 1201;
Best Local Similarity 39.9%; Pred. No. 25;
Matches 63; Conservative 25; Mismatches 70; Indels 0; Gaps 0;

QY 154 GCCATTTCACCCGATCAAGCTTTCGACAGATGATGATGAACTTACCTT 213
DB 943 GACTTCTCMGCGCCGCGGCGGCTGTGTGCGGCGCGCCCTTCGCGGT 1002
QY 214 GGTCTCATTCGCGTGGATTAACATGATGCGCATCTTCTTGGCGATGACACTT 273
DB 1003 GSTKSGGTTSMGGCCMCTTKAMAGTCGGGCGVGHYTTTGTGTTCCCMGCCCTC 1062
QY 274 TTGAGCGTTGCGCGCTGACCATCGTTTGTGACT 311
DB 1063 CAGTATGTCGAGGAGTACATCTCTGTATAGT 1100

RESULT 11

CB239541

LOCUS

DEFINITION

CB239541 413 bp mRNA linear EST 30-APR-2003
RSH13B11 two-month-old roots from clone 'Beaupre' grown for 19 days
under restricted irrigation Populus balsamifera subsp. trichocarpa

ACCESSION

CB239541

GI:30233030

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 413)
Kohler, A., Delanuelle, C., Martin, D. and Martin, F.
The poplar root transcriptome: analysis of 7000 expressed sequence
tags

JOURNAL

Unpublished (2003)
Contact: Martin FM

COMMENT

Equipe de Microbiologie Forestiere
Institut National de la Recherche Agronomique
Centre INRA de Nancy, 54280 Champenoux, France
Tel: +33 383 39 40 80
Fax: +33 383 39 40 69
Email: fmarlin@nancy.inra.fr
Insert Length: 413 Std Error: 0.00
Seq primer: Fornat 5' AAGCGCGCATGTGTGTATACC.

FEATURES

source

1..413
/organism="Populus balsamifera subsp. trichocarpa x
Populus deltoides"
/mol_type="mRNA"
/cultivar="Beaupre"
/db_xref="taxon:3695"
/def_stages="two-month-old"
/clone_lib="two-month-old roots from clone 'Beaupre' grown
for 19 days under restricted irrigation"
/note="Organ: root; Vector: pTriplex2; cDNA library of
roots from two-month-old Populus trichocarpa Torr. & Gray
x deltoides Bartt. Ex Marshall (clone 'Beaupre') grown for
19 days under restricted irrigation to reach 50% of the
transpiration rate of fully watered plants. The cDNA
library was constructed from 1 ug of total RNA using the
SMART cDNA synthesis kit (Clontech, Palo Alto, CA, USA)
according to the manufacturer's instructions. The
resulting cDNA was packed into lambda phages using the

ORIGIN

Gigapack III Gold packaging kit (Stratagene, La Jolla,
CA). The pTriplex2 phagemid clones in Escherichia coli
were obtained by using the mass in vivo excision protocol
according to the manufacturer's instructions (Clontech)."

Query Match 10.7%; Score 34.6; DB 14; Length 413;
Best Local Similarity 60.4%; Pred. No. 19;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 65 TCCGGCGGTCCTGCTTTTATCTTAAGCCCTTACTGATCAATTTGTGGCAAA 124
DB 71 TCCAGGGCTCCCTCTGAGACTATTAACCCCTCCAGGATTCCTTCAGGTCAAAA 130
QY 125 TGGGATGTGATGCCAGCAGGAATCTTGC 155
DB 131 TGGGATTCGGGCGCACAGGATTCCTGCG 161

RESULT 12

CK297477

LOCUS

DEFINITION

CK297477 853 bp mRNA linear EST 15-DEC-2003
EST1760191 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NEMD912 5'
end, mRNA sequence.

ACCESSION

CK297477

GI:39883902

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Staskiewicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Other ESTs: EST760192
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-arrays@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA TAG.

JOURNAL

Unpublished (2003)
Contact: Robin Buell

COMMENT

The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-arrays@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA TAG.

FEATURES

source

1..853
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMD912"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Pseudomonas syringae pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 853;
Best Local Similarity 47.5%; Pred. No. 25;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Fukuda, S., Hashizume, M., Hayashida, K., Hirozane, T., Hori, F., Imokuni, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, N., and Hayashizaki, Y. Direct Substitution

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source
1..728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F830034E14"
/tissue_type="activated spleen"
/clone_lib="RIKEN full-length enriched, activated spleen"

ORIGIN

Query Match 10.7%; Score 34.8; DB 13; Length 728;

Best Local Similarity 55.9%; Pred. No. 20;

Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 188 CGATGATCTGAAGACTCTTAACCTTGTCTCATTCGCCGTGGATTACAGTGTGGCC 247

DB 244 CGACGTAGCTGATCTTTAATCCCTTGCTAGAGCTGCTGTTGTAAGAGAGCTGTGC 303

QY 248 ATCTTCTTGCGGCTGACGACCTTGTGAGCGTGGCGCTGGACCATCTTTTGT 305

DB 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

QY 304 AACATCTGTCTGCGACGCCCTCTTCTTGCTGTGCTGTGTAAGGCAATGCTTTCTGT 361

FEATURES

source
1..808
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064373"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library." Location/Qualifiers

ORIGIN

Query Match 10.7%; Score 34.8; DB 12; Length 808;

Best Local Similarity 53.7%; Pred. No. 21;

Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 133 TGGATGCCAGCAGGAGATCTTGCATTTTGACCCGATCAAGTTCCAGCAATGCGATA 192

DB 523 TGGAGGCTCCATGTTCCAACTTAATGTGAGGCAAGCTGGAGCATCAACAGACAGAT 582

QY 193 GATCTGAAGACTCTTAACCTTGTCTCATTCGCCGTGGATTACAGTGTGGCCACTT 252

DB 583 GTTCAGAGATGAGAGAGCTTGGGTTCAAAAGCTTGTGTGCTACTGTAGATGCCCTGTA 642

QY 253 CTTGGCGGCTGCAGC 266

DB 643 CTTGGCAATAGCG 656

RESULT 10

BX387771

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

11201 bp mRNA linear EST 08-MAY-2003 BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

11201 bp mRNA linear EST 08-MAY-2003 BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

11201 bp mRNA linear EST 08-MAY-2003 BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

11201 bp mRNA linear EST 08-MAY-2003 BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

11201 bp mRNA linear EST 08-MAY-2003 BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA530007"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

Query Match 10.9%; Score 35.2; DB 28; Length 910;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 192 AGATCTGAAGACTTAACTTGTGCTCATTGCGCTGACATGAGTGTGGCGGAT 251
Db 272 AAATCTACGCGCTAAATCTCTTCCCAATCCGACATCGGACGCGCGCGCC 331

Qy 252 TCTTGGCGGTCGACGACCTGTGAGCGTGGCGCTGACCATGCTTTT 303
Db 332 TGTGGCGGCGGATCTCGCTGGGACGACCGGATGCTGCGCGCTTT 383

RESULT 7 1101 bp DNA linear GSS 04-JUN-1999
CNS00BD8 Drosophila melanogaster genome survey sequence TERT end of BAC #
LOCUS BACR23H11 of RPCR-98 library from Drosophila melanogaster (fruit
DEFINITION fly), genomic survey sequence.
ACCESSION AL056828.1 GI:4937496
VERSION AL056828
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Ogoe and
Aaron Mammoe in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCR-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..1101

SOURCE
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR23H11"
/clone_1db="RPCR-98"
/note="end : TERT3"

ORIGIN
Query Match 10.9%; Score 35.2; DB 29; Length 1101;
Best Local Similarity 24.4%; Pred. No. 18;
Matches 62; Conservative 81; Mismatches 111; Indels 0; Gaps 0;

Qy 29 TTGTTGGCAGTATGTCAGTACTTTTGGCGTCCGGCGGCTCCGTTCTTATTC 88
Db 847 TTGSGKTGTTGSGAASDTGTTATATTSTTTTBSKSTSTBSKMAATKWSASTTTT 906

Qy 89 TTAAGCCCTTACGTCATCAATTTGTGGGCAAAATGCGATGTGATGCCACAGAA 148
Db 907 WSGASGTSITKTBKSGTSTRTVTTTWSGSAVAASGTTTWTSTWSTBKABBA 966
Qy 149 TCGTGCATTTTGGACCGCATCAAGCTTTCGACCAATGCGATGATCTGAAGACTTAA 208
Db 967 ATSAHTABSDPTAABSKTSBSBSSBDBGGTTPASGSAATTTKTTTITTTTWW 1026
Qy 209 CTTTGTCTCATGTCCTGCGATTCATGATGTCGCGCATCTTTTGGCGGTGACGCA 268
Db 1027 TATTTNSBSWBXATTTTDTAAATTAATVSTNSNSNNNTTBSRPAATPAAVSA 1086
Qy 269 CTTTGTGAGGCTT 282
Db 1087 GTBTBBSVATB 1100

RESULT 8 728 bp mRNA linear EST 17-DEC-2002
BY765711
LOCUS BY765711 RIKEN full-length enriched, activated spleen Mus musculus
DEFINITION cDNA clone F830034E14 3', mRNA sequence.
ACCESSION BY765711
VERSION BY765711.1 GI:27203910
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 728)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamana, A.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schiraldi, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bristol, V.,
Chochia, C.F., Corbani, L.E., Cousins, S., Daille, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Gilmord, S.,
Gustindich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kongave, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglocz, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Warande, Y.,
Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yangisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kitahara, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@ac.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

LOCUS CF338060 475 bp mRNA linear EST 18-AUG-2003
 DEFINITION UMT--08-M05.b1 AtUMT-overexpressing transgenic rice plasmid cDNA
 library (UMT) Oryza sativa cDNA clone UMT--08-M05, mRNA sequence.
 ACCESSION CF338060
 VERSION CF338060.1 GI:33824491
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 475)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 JOURNAL Contact: Nahm B.H.
 COMMENT Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
 Location/Qualifiers
 source
 1..475
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="UMT--08-M05"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="AtUMT-overexpressing transgenic rice plasmid
 cDNA library (UMT)"
 /note="Vector: PCR4-TOPO, Site_1: EcoRI; Oligo-capped mRNA
 was reverse transcribed and then used for PCR. mRNA was
 prepared from Arabidopsis thaliana cDNA library
 methyltransferase overexpression line."

ORIGIN

Query Match 10.9%; Score 35.2; DB 14; Length 475;
 Best Local Similarity 50.6%; Pred. No. 13;
 Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 51 CATTACTTTGCGCTCGGCGGCTCCGTTCTTAATCCTTAAGCCCTCAAGTAACACA 110
 DB 255 CTTCTTCTTGCGCTCGGCGGCTCCGTTCTTAATCCTTAAGCCCTCAAGTAACACA 196
 QY 111 ATTGTGGGCAAAATGGCATGTGATGCCAGAGAAATCTTGCATTTTGAACGCATC 170
 DB 195 CGGCGTGTGCGCGTCTCACTGAGCTTGGCAGCGGCTCTTGGCAAGTCGCGCAGTTT 136
 QY 171 AACGTTGCGAGCATGCATGATCTGAAGACTTAACCTTGTGCT 218
 DB 135 CACGCGCTTCAAGAGCCATTTCTGATGCTTCTGTCTGTCT 88

RESULT 5
 BZ720836 786 bp DNA linear GSS 24-FEB-2003
 LOCUS PUCF0505TD ZM 0.6_1.0 KB Zea mays genomic clone ZMBRta137A09,
 DEFINITION genomic survey sequence.
 ACCESSION BZ720836
 VERSION BZ720836.1 GI:28512409
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 786)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

RESNICK,A., FRASER,C.M., YUAN,Y., SAN MIGUEL,P., MA,J. and
 BENNETZEN,J.
 Maize Genomics Consortium
 Unpublished (2003)
 JOURNAL Contact: Cathy WhiteLaw
 COMMENT TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 source
 1..786
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMBRta137A09"
 /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 cot selected genomic DNA library"

ORIGIN

Query Match 10.9%; Score 35.2; DB 28; Length 786;
 Best Local Similarity 53.7%; Pred. No. 16;
 Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 189 CATGATCTGAAGCTTAACCTTGTGCTTCATTTGCCGTTGACATGCTGGCGCA 248
 DB 480 GAAGCAATCAAAATGTTGTCTTACCTCAAGGTGGCCAACTGTGTGCTTA 539
 QY 249 TCTTCTGGCGCTGACGACACCTTGTGAGGCTTGGCGTGGCACCATCGTTTGTGG 308
 DB 540 TCATTAGTGTGATCTTAAGATGTTATCATTTGCGCCCAACAATCTTGTGAGTCGG 599
 QY 309 ACTGCTGAATTTTTC 324
 DB 600 TATGTTGCAATGTGC 615

RESULT 6
 CC362201 910 bp DNA linear GSS 16-MAY-2003
 LOCUS PUHTK88TD ZM 0.6_1.0 KB Zea mays genomic clone ZMBRta130007,
 DEFINITION genomic survey sequence.
 ACCESSION CC362201
 VERSION CC362201.1 GI:30831601
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 910)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 JOURNAL Other GSSs: PUHTK88TB
 COMMENT Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 source
 1..910
 /organism="Zea mays"
 /mol_type="genomic DNA"

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_pac.htm.

FEATURES

Location/Qualifiers
1..995
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12G04"
/clone_id="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 11.7%; Score 37.8; DB 29; Length 995;
Best Local Similarity 34.0%; Pred. No. 2.9;
Matches 54; Conservative 41; Mismatches 64; Indels 0; Gaps 0;

QY 164 CCGCATCAAGCTTTGCGAGCAATCGATGATCTGAAGCTTAACTTTGTTCTGATG 223
DB 872 CAGTMTTACGACCGGSGTGTGMAAGSASASATGVRSTCGCTTTTMTTAAAT 813
QY 224 CCGTTGGATTACAGTGTGCGGATCTTCTGGCGGTCGACGACCTTTGAGCGTGG 283
DB 812 CCKTTSBTWKAACSCMTTSSGSMASWTGCTGKKGAGSAGAAAGGCGCAKKS 753
QY 284 GCGCTGCACCATGCTTTTGTGACTGTGAATCTT 322
DB 752 SSGTGTGKXTSARKKTGTGTGSGTGSDTATCTTT 714

RESULT 2 1353 bp DNA linear GSS 17-DEC-2002
LOCUS BZ557685
DEFINITION pac61-60.663.g1 pac61-60 Pseudomonas aeruginosa genomic clone
ACCESSION BZ557685
VERSION BZ557685.1 GI:27170831
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1353)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V. Whole-Genome-Sequence Variation among multiple isolates of *Pseudomonas aeruginosa* library J. Bacteriol. (2002) In press
JOURNAL Contact: Chris K. Raymond
COMMENT Genome Center
University of Washington
Box 353145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: shotgun.

FEATURES

Location/Qualifiers
1..1353
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pac61-60.663"
/clone_id="pac61-60"
/note="Clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

1..1353
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pac61-60.663"
/clone_id="pac61-60"
/note="Clinical isolate 1-60 Whole genomic shotgun library."

Query Match 11.0%; Score 35.8; DB 28; Length 1353;
Best Local Similarity 57.7%; Pred. No. 13;
Matches 64; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTCATTTGCGGATTAAGTGTGGCGCATCTTGGCGGTG 263
DB 86 TATAGCCCTTCTCGTCGGGCTTGAAGTGCCTGGCGCGCTTATCTGCTGATCC 145
QY 264 ACGACCTTGTGACGCTTGGCGCTGACACATGCTTTTGTGCACTGG 314
DB 146 GCGCAGGCGCTTCTCATGCGGSGTGTGCGGCTTGTGCGGCGCAGGT 196

RESULT 3

CF302597/c 471 bp mRNA linear EST 15-AUG-2003
LOCUS 7LEAF--08-E15.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION sativa cDNA clone 7LEAF--08-E15, mRNA sequence.
ACCESSION CF302597
VERSION CF302597.1 GI:33674358
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 471)
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H. Large-scale Sequencing Analysis of Rice ESTs Unpublished (2003)
CONTACT: Nahm, B.H.
GENOMICS AND GENETICS INSTITUTE, GREENGENE BIOTECH INC., DIVISION OF BIOENGINEERING AND BIOINFORMATICS, WYONGI UNIVERSITY YONGIN, KYONGGI, KOREA
TEL: 82 31 330 6193
FAX: 82 31 321 6355
EMAIL: bhnam@bio.com, bhnam@bio.myongji.ac.kr.

FEATURES

Location/Qualifiers
1..471
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF--08-E15"
/issue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_id="Rice leaf plasmid cDNA library II (7LEAF)"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; RNA was capped with oligoribonucleotides and then used as templates for RT-PCR."

ORIGIN

Query Match 10.9%; Score 35.2; DB 14; Length 471;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 51 CATTAATTTTGGGCTCCGGGCGGCTTCCGTTCTTAATCCCTAGCCCTAAGGATACCA 110
DB 224 CTTCTCTTGGGCTCCGGGCGGCTTCCGTTCTTAATCCCTAGCCCTTGGGCGGCGG 165
QY 111 ATTGTGGCAAAATGCGATGTGATGCGACAGAGATCTTGCATTTTGAACGCAATC 170
DB 164 CCGGCTTGTGCGGCTGTACTAGGCTGCGACCGGCTCTTGGCAAGTCCGCAAGTTT 105
QY 171 AACGTTTCCAGCAATGCATGATCTGAAGACTCTTAACCTTTGGCT 218
DB 104 CACGGCTTCAAGAGACCATCTTCTGATGCTTCTGCTGTCT 57

RESULT 4

CF338060/c

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 1049.18 Seconds

(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324
Sequence: 1 atgacacacgattcttccttcg.....ttgacgtgtaaccttcc 324

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estda:*
2: em_estnum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pio:*
25: em_gss_rtd:*
26: em_gss_pbg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	11.7	995	29	CNS005NG
C 2	35.8	11.0	1353	28	BZ557685
C 3	35.2	10.9	471	14	CF302597
C 4	35.2	10.9	475	14	CF338060

Result No.	Score	Query Match	Length	ID	Description
C 5	35.2	10.9	786	28	BZ720836
C 6	35.2	10.9	910	28	CC362201
C 7	35.2	10.9	1101	29	CNS005NG
C 8	34.8	10.7	728	13	BY765711
C 9	34.8	10.7	808	13	BY765711
C 10	34.8	10.7	1201	13	BY765711
C 11	34.6	10.7	413	13	CB239541
C 12	34.6	10.7	853	14	CK297477
C 13	34.6	10.7	890	14	CK288648
C 14	34.6	10.7	922	14	CK295854
C 15	34.6	10.7	936	14	CK295424
C 16	34.6	10.7	951	14	CK292130
C 17	34.4	10.6	447	29	CNS002NHL
C 18	34.4	10.6	1201	13	BY765711
C 19	34.2	10.6	1062	28	BZ557686
C 20	34.2	10.5	423	28	B44402
C 21	33.8	10.4	1088	29	CNS046RO
C 22	33.6	10.4	833	13	BY765711
C 23	33.4	10.3	580	12	BM309296
C 24	33.2	10.2	478	14	H65669
C 25	33.2	10.2	586	14	CD091362
C 26	33.2	10.2	859	10	BF341287
C 27	33.2	10.2	1044	28	BZ553873
C 28	33.2	10.2	569	12	BM100089
C 29	33.2	10.2	720	29	CG04816
C 30	33.2	10.2	761	29	CG04816
C 31	33.2	10.2	796	29	CG04816
C 32	33.2	10.2	818	28	CC005623
C 33	33.2	10.2	958	29	CC005623
C 34	32.8	10.1	390	10	AM456947
C 35	32.8	10.1	399	10	AM491785
C 36	32.8	10.1	416	10	AM457742
C 37	32.8	10.1	463	10	AM489067
C 38	32.8	10.1	469	10	BF460793
C 39	32.8	10.1	486	12	BI135202
C 40	32.8	10.1	496	12	BI134847
C 41	32.8	10.1	536	12	BI901653
C 42	32.8	10.1	565	12	BI901912
C 43	32.8	10.1	602	10	BF207637
C 44	32.8	10.1	638	14	CF582833
C 45	32.8	10.1	3361	11	AK083916

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CNS005NG 995 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TERT end of BAC #
BACR12604 of RPI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL060428.1 GI:4943359
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 995)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Oosawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of

QY 178 CGGAGCAATGCGATTAATCTGAAGACTCTTAACCTTTGCTCATGCGCTGGATTACA 237
DB 15721 GGGTCCACTCTCTGGGGAAGAGGCGCTTGAAGACAGACCAATCTGCTGTGGTGTGA 15780
QY 238 GTGTGGCGCATCTTCTTGGCG 259
DB 15781 GTGTACTCTATCTCTCTGGTG 15802

RESULT 14
US-10-796-280-12378
; Sequence 12378, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12378
; LENGTH: 234309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(234309)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-796-280-12378

Query Match 8.8%; Score 28.4; DB 6; Length 234309;
Best Local Similarity 52.5%; Pred. No. 81;
Matches 62; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 206 TAACTTGTGTCATGTCCTGCGCTTGCATTAAGTGTGGCGCATCTTCTTGGCGGTGAC 265
DB 62038 TCACCTTAGGCGCTTTAAGCCTGAAGAGCGTGGCGCCTTACATCTGCTTTTGAATTTC 62097
QY 266 GCACCTTGTGAGCGTGGCGCTGACCACTGCTTTTGTGACCTGTGAATCTTTT 323
DB 62098 GTCCTTGTGTGAAGTGGCTAAGGCCCTTCCGCTCTGTGTGAAGAGGGGTTTTT 62155

RESULT 15
US-10-767-701-27384/C
; Sequence 27384, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 27384
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 6858297
US-10-767-701-27384

Query Match 8.7%; Score 28.2; DB 6; Length 490;
Best Local Similarity 51.2%; Pred. No. 5.8;
Matches 66; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 192 AGATCTGAAGACTTAACCTTGGTCTCATTTGGCGGTGACATTACAGTGTGGCGCATCT 251
DB 282 AGGTGTACCAAGTGTAGCACTTGTGTAGCTGCCAAGGACGAGTGTGAGCCAGCC 223

QY 252 TCTTGGCGGTGACGACCACTTGTGAGGCTGGCGTGGCAACATCGTTTGTGGACT 311
DB 222 AGTTGGCGGTGACGAGGTGACGACCTGGTGGCGGTGGCAACACAGACTTGTGGACT 163
QY 312 GGTGAATCT 320
DB 162 GGGCAACT 154

Search completed: April 17, 2004, 20:13:04
Job time : 39.4407 secs

QY 143 CAGGATCTTGGCATTGTAACCGCATCAACGTTTCCGACGATGCAATGATCTGAGA 202
 DB 284 CGGGTATCTGGCGGATATACCGACGACTATTCACCGCAGAGGTTTGGCCGACGG 225
 QY 203 CTCRAACCTTGGCTCATTTGGCGGTGCGATTACAGTGTGGCGCATCTTCTTGGCGGTC 262
 DB 224 CACTTAACCGTACGGGATTTGGCTTCCGACGACTGATGGCTTCAACATGGGCGATG 165
 QY 263 GACGACCTTGTGTG 276
 DB 164 TTCTGATCATGATG 151

RESULT 10
 US-10-108-605A-248/c
 ; Sequence 248, Application US/10108605A
 ; GENERAL INFORMATION:
 ; APPLICANT: Broadus, Julie
 ; APPLICANT: Stam, Lynn
 ; APPLICANT: Bachmann, Jane
 ; APPLICANT: Kamdar, Kim
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 ; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 ; FILE REFERENCE: 31133B
 ; CURRENT APPLICATION NUMBER: US/10/108, 605A
 ; PRIOR FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US 09/761,142
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/176,418
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 248
 ; LENGTH: 8577
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-10-108-605A-248

Query Match 8.8%; Score 28.4; DB 6; Length 8577;
 Best Local Similarity 60.3%; Pred. No. 18;
 Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
 QY 209 CTTTGTCTCATTTCCGTTGCGATTACAGTGTGGCGCATCTTCTTGGCGGTGAGCA 268
 DB 83 CCGTGTGTGTGTGCGCGCACCTGCTACTGTGTGCTTGTGCTTGGCGAGGACGAC 24
 QY 269 CTTTGTGAGCGTTGGCG 286
 DB 23 CCGCTTCAACGCGGTG 6

RESULT 11
 US-10-796-280-12417/c
 ; Sequence 12417, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12417
 ; LENGTH: 20618
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12417

Query Match 8.8%; Score 28.4; DB 6; Length 20618;
 Best Local Similarity 51.6%; Pred. No. 27;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGCGGTCCTCGTTCTTAATCCTTAAGCCCTACGATGATACAAATTTGGGCAAA 124
 DB 8143 TCCATCATTTCCACTGTATTAACCTTTGGCTCAAGCAGATACCTCTTGTGATTA 8084
 QY 125 TGGCAGATGATGACGACGAGAAATCCTTGCCATTTTGACCGCATCAACGTTTGGACGA 184
 DB 8083 TGCTGTCTCACTGTTTCATCTCCGCGCCCTGCGCCCTGTCTTCCCAACTATTTCCACCA 8024
 QY 185 ATGCA 190
 DB 8023 AGGCA 8018

RESULT 12
 US-10-796-280-12418
 ; Sequence 12418, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12418
 ; LENGTH: 46075
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12418

Query Match 8.8%; Score 28.4; DB 6; Length 46075;
 Best Local Similarity 51.6%; Pred. No. 39;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 65 TCCGGCGGTCCTCGTTCTTAATCCTTAAGCCCTACGATGATACAAATTTGGGCAAA 124
 DB 3888 TCCATCATTTCCACTGTATTAACCTTTGGCTCAAGCAGATACCTCTTGTGATTA 3947
 QY 125 TGGCAGATGATGACGACGAGAAATCCTTGCCATTTTGACCGCATCAACGTTTGGACGA 184
 DB 3948 TGCTGTCTCACTGTTTCATCTCCGCGCCCTGCGCCCTGTCTTCCCAACTATTTCCACCA 4007
 QY 185 ATGCA 190
 DB 4008 AGGCA 4013

RESULT 13
 US-10-796-280-12452
 ; Sequence 12452, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12452
 ; LENGTH: 69359
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12452

Query Match 8.8%; Score 28.4; DB 6; Length 69359;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 118 GGGAAATGGCATGTGATGACGAGGATCTTGCATTTTGAACGATCAACGTTT 177
 DB 15661 GGCAACTGGGATTTAGAACTAGGCACTTGACTCAGAGCTGTGCCCCCTGTGCATT 15720

Query Match	8.8%	Score 28.4	DB 1	Length 500
Best Local Similarity	50.7%	Pred. No. 4.9		
Matches 68	Conservative 0	Mismatches 66	Indels 0	Gaps 0

US-10-789-400-36
Sequence 36, Application US/10789400
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
APPLICANT: HUMAN SERVICES
APPLICANT: COLLINS, Peter L.
APPLICANT: Biacchesi, Stephane
APPLICANT: Buchholz, Ursula
APPLICANT: Skidopoulos, Mario H.
APPLICANT: Murphy, Brian R.
TITLE OF INVENTION: RECOMBINANT HUMAN METAPNEUMOVIRUS AND ITS USE
FILE REFERENCE: 4239-67783
CURRENT APPLICATION NUMBER: US/10/789, 400
CURRENT FILING DATE: 2004-02-27
PRIOR APPLICATION NUMBER: US 60/451,119
PRIOR FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/478,667
PRIOR FILING DATE: 2003-06-13
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.2
SEQ ID NO 36
LENGTH: 13350
TYPE: DNA
ORGANISM: human metapneumovirus
US-10-789-400-36

Query Match
Best Local Similarity 9.4%; Score 30.4; DB 6; Length 13350;
Pred. No. 4.5;
Matches 79; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 14 TCTCTGTATCTCTCTGTGTGCGAGTATGCGATCTTCTTCCCTCCGGCGG 73
DB 1687 TCTCAGATATGAAGAAGAAAGATGCAATCTTCACTTGAAGAAAGATA 1746
QY 74 TTCGTTCTTATCTTATAGCCCTAGCATGATGAGCAATTTGGGCAAAATGGCATG 133
DB 1747 CTTATCTTATAGCATGAGCCGAGATTTGAAATGAAGAAATTAAGATATAT 1806
QY 134 GGATGCCAGCAGAAATCTTGCATTTTGAACCGCATAC 173
DB 1807 TAGGGCTATTAAGAACTCAACATTTGCTACAGCAGACC 1846

RESULT 3
US-10-767-471-10749/c
Sequence 10749, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10749
LENGTH: 35895
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(35895)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10749

Query Match
Best Local Similarity 9.3%; Score 30.2; DB 6; Length 35895;
Pred. No. 8.2; Mismatches 58; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 201 GACTCTAATCTTGTCTCATTTGCGGTTGCGATTACAGTGTGGCGCATCTTCTGGCGG 260
DB 12662 GCGCTCAGCCAGCTGCTCTCTTGGGTTGTCTTGGCAGCAGAGCCATCTCTGGGCTG 12603

QY 261 TCGACGACCTTTGTGAGCGTTGGCGCTGGCAGCACCATGTTTGTGAGCTGTGAATCT 320
DB 12602 CCCACACATCTGCTTGGCGCTTCCCTAGCTCCAGACAGTGAATGTTTGAACCTAGACCT 12543
QY 321 TTT 323
DB 12542 TCT 12540

RESULT 4
US-10-767-701-5423/c
Sequence 5423, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5423
LENGTH: 731
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28WAY03-CLUS132149_1
US-10-767-701-5423

Query Match
Best Local Similarity 8.9%; Score 28.8; DB 6; Length 731;
Pred. No. 4.3; Mismatches 52; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 149 TCCITGCCATTTTGAACCGCATCAAGCTTTCGACGAATGAGATGATGAAGCTTAA 208
DB 240 TCTCGCCAGTTTGTCTCGGACACACATCCACCATGGAATGGTTGACGATTAACATG 181
QY 209 CCTTGGCTCATTTGCCCTTGGCATTAAGTGTGGCGCATCTTCTGGCGG 260
DB 180 CATTCATCATCTTGAAGTGTGAGTGAAGCTGAACCACTCTCGGCGTG 129

RESULT 5
US-10-767-471-10669/c
Sequence 10669, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10669
LENGTH: 158027
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(158027)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10669

Query Match
Best Local Similarity 8.9%; Score 28.8; DB 6; Length 158027;
Pred. No. 50; Mismatches 44; Indels 0; Gaps 0;
Matches 54; Conservative 2; Mismatches 44; Indels 0; Gaps 0;
QY 48 AGTCATTAATCTTGTGCGCTCGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATC 107
DB 48997 AGCTTTGAATCTTAAACCAAGCTGATGATGTAATCTGTCTCCCAAAATTC 48938

Tue Apr 20 06:47:26 2004

us-10-608-504-6_copy_853_1176.rn

Page 7

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 60315015
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEO ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-476-900A-4

	Query Match	Best Local Similarity	Score 29.8	DB 3	Length 5110
	Matches 52	Conservative 0	Pred. No. 3	Indels 0	Gaps 0
Qy	118	GGCAAAATGGGCAATGGATGGCCACAGAAATCCTTGACATTTGACCGCATCAAGCTTT	177		
Db	4500	GGCACTTGGCGCATGTCATGCGCATGATGATGATAGCGCCTTCAAAATATCCACGATT	4441		
Qy	178	CGCAGCAATGGCATGATCTGAAGACTCT	206		
Db	4440	CGGAAAAAGCGAGAGAGAGAGAGGATTT	4412		

Search completed: April 19, 2004, 18:38:22
Job time : 28.2807 secs

REFERENCE/DOCKET NUMBER: BYLR-0003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-404-531B-27

Query Match 9.2%; Score 29.8; DB 2; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 118 GCGAAATGGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 177
DB 4025 GGCAGCTTGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 3966

QY 178 CGCAGCAATGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 206
DB 3965 CGGAAAAGCGCGAGAGAGAAATGATTT 3937

RESULT 11
US-08-476-900A-27/C
Sequence 27, Application US/08476900A
Patent No. 6031150
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
Patent No. 6031150
TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESS: No. 60311501st
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-476-900A-27

Query Match 9.2%; Score 29.8; DB 3; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 118 GCGAAATGGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 177
DB 4025 GGCAGCTTGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 3966

QY 178 CGCAGCAATGCGATGTGATGCGAGAGAAATCCTTGGCATTCGACGATCAAGCTT 206
DB 3965 CGGAAAAGCGCGAGAGAGAAATGATTT 3937

RESULT 12
US-08-488-546A-27/C
Sequence 27, Application US/08488546A
Patent No. 6054313
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 6054313
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESS: No. 60543131st
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,546A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,531
FILING DATE: 15-MARCH-1995
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533

RESULT 8
US-08-809-254A-4/C
Sequence 4, Application US/08809254A
Patent No. 6660852
GENERAL INFORMATION:
APPLICANT: KESHI ET AL
TITLE OF INVENTION: PROBE FOR DIAGNOSING INFECTIOUS DISEASES
FILE REFERENCE: 19036/33/67
CURRENT APPLICATION NUMBER: US/08/809,254A
CURRENT FILING DATE: 1997-05-16
PRIORITY FILING DATE: 1995-10-02
PRIORITY APPLICATION NUMBER: JP 236348
PRIORITY FILING DATE: 1994-09-30
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 5829
TYPE: DNA
ORGANISM: Artificial sequence
FEATURES:
OTHER INFORMATION: Synthetic probe
US-08-809-254A-4

Query Match 9.5%; Score 30.8; DB 4; Length 5829;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 122 AATGGCGATGTGATGCGCAGAGATCCTTGCCATTGACCGCATCAAGTTGCCA 181
DB 671 ATATGGCGAGAAATTCATCAGTTATCTTGCCACCTTCTCGCATATTTAAAGGA 612
QY 182 GAATGGCATAGCTCTGAAGCTCAACCTTGGCTC 219
DB 611 AAGATGCTAGAGATATGGAATATCGCTCTTTGTATC 574

RESULT 9
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMCU

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZpT-fls
US-08-232-463-14

Query Match 9.4%; Score 30.6; DB 1; Length 7218;
Best Local Similarity 1.0%; Pred. No. 1.9;
Matches 3; Conservative 179; Mismatches 133; Indels 0; Gaps 0;

QY 10 GATTCTCCTGATCTCTCTGTTGTGCGCATGATGCGATCTTACTTTGCGCTCGG 69
DB 1064 GATYYY 1123
QY 70 GCGTTCGCTCTTAATCCCTAAGCCCTAGCATGATGCAATTTGGCGAAATGSCG 129
DB 1124 YY 1183
QY 130 ATGTGATGCGCAGAGATCCTTGCAATTTGACCGCATCAAGCTTGGCGCATGCG 189
DB 1184 YY 1243
QY 190 ATAGATCTGAAGACCTCTAACCTTGCTCATGTCGCTGAGATGAGTGGCGCAT 249
DB 1244 YY 1303
QY 250 CTCTTGCGGCTGAGCAGCACTTGTTGAGCGTTGCGCTGGCAGCATGTTTGTGA 309
DB 1304 YY 1363
QY 310 CTGTCGATCTTTC 324
DB 1364 YYYYYYYYYYYYYY 1378

RESULT 10
US-08-404-531B-27/C
Sequence 27, Application US/08404531B
Patent No. 5863724
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 5863724
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5863724is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,531B
FILING DATE: 15-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13729
LENGTH: 594
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13729

Query Match
Best Local Similarity 56.8%; Pred. No. 0.023;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTTCATTCGCTGCGCATTTACAGTGTGGGCGCATCTTTGGGGGTG 263
DB 407 TATAGGCTTGGCCCTGCGTGGAGTGAAGTCCGCGGGCGGCTTGATCTTGCTGATCC 348
QY 264 AGGCACCTTGTGAGGCGTGGCGGCACATCGTTTGTGAGCTGCT 314
DB 347 GCGCCAGGGGCTTCGTGATCGGGGTCTGCTCGCGCTTGCTGCGGCAAGT 297

RESULT 5
US-09-252-991A-13832/C
Sequence 13832, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13832
LENGTH: 678
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13832

Query Match
Best Local Similarity 56.8%; Pred. No. 0.025;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTTCATTCGCTGCGCATTTACAGTGTGGGCGCATCTTTGGGGGTG 263
DB 472 TATAGGCTTGGCCCTGCGTGGAGTGAAGTCCGCGGGCGGCTTGATCTTGCTGATCC 413
QY 264 AGGCACCTTGTGAGGCGTGGCGGCACATCGTTTGTGAGCTGCT 314
DB 412 GCGCCAGGGGCTTCGTGATCGGGGTCTGCTCGCGCTTGCTGCGGCAAGT 362

RESULT 6
US-09-328-352-1339
Sequence 1339, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 1339
LENGTH: 1404
TYPE: DNA
ORGANISM: Acinetobacter baumannii

US-09-328-352-1339

Query Match
Best Local Similarity 55.5%; Pred. No. 0.32;
Matches 61; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 136 ATGCCAGGAGATCCCTGCAATTTGACCGCATCAAGCTTGGCAGATGATAGAT 195
DB 541 ATTCCATGGAGATATGGCATATTATGACATTAAGTCAATGCCAAATGAAAGAG 600
QY 196 CTGAAGACTTAACCTTGTCTCATTCGCGGTGCGATTACAGTGTGCG 245
DB 601 CTTACCGTTAACTTTGATCTTATTTGAGATTGTGCTCTTACGTCGCG 650

RESULT 7
US-08-961-527-56
Sequence 56, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 391
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 2387 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-56

Query Match
Best Local Similarity 52.7%; Pred. No. 0.51;
Matches 68; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 112 TTTGTGGCAAAAGGCGATGTGATGCCAGAGAAATCCCTTGCAATTTGACCGCATCA 171
DB 962 TTTGTGGAAAAAGACAATTAATTTGACAGAAAGCAATTTGAAAAATGATTTTA 1021
QY 172 AGCTTGGCAGCAATGATGATCTGAAGCTTAACTTAACTTGTGCTCATTCGCGTGGC 231
DB 1022 GCAATTCAGCGTATTTCTCAATGAGAAATTAACAGTTAGTCTTCTGCTTAAT 1081
QY 232 ATTAAGTG 240
DB 1082 ATGATGTG 1090

QY 241 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 300
DB 241 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 300
QY 301 TTGTGTGACTGTGATCTTTTC 324
DB 301 TTGTGTGACTGTGATCTTTTC 324

RESULT 2

US-09-471-803A-1
Sequence 1, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471,803A
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnf
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brne
OTHER INFORMATION: ATCC14752
US-09-471-803A-1

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAACTGATTTCTCTGTAATCTCTTGTGTGCGAGATGTGAGTCATTACTTTT 60
DB 853 ATGCAACTGATTTCTCTGTAATCTCTTGTGTGCGAGATGTGAGTCATTACTTTT 912
QY 61 GCGCTCCGGGGGGGTTCCGTTCTTAATCCTTAAGCCCTTACGTGATCACAATTTGTGGC 120
DB 913 GCGCTCCGGGGGGGTTCCGTTCTTAATCCTTAAGCCCTTACGTGATCACAATTTGTGGC 972
QY 121 AAAATGGCGATGTGATGCGACGAGAGATCTTGGCCATTGTGACCGCATCAAGCTTTCG 180
DB 973 AAAATGGCGATGTGATGCGACGAGAGATCTTGGCCATTGTGACCGCATCAAGCTTTCG 1032
QY 181 AGCAATGCGATGATGATGAGATCTTAACCTTGTGTCATGTCGCTGTGAGTTACAGTG 240
DB 1033 AGCAATGCGATGATGATGAGATCTTAACCTTGTGTCATGTCGCTGTGAGTTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 1152
QY 301 TTGTGTGACTGTGATCTTTTC 324
DB 1153 TTGTGTGACTGTGATCTTTTC 1176

RESULT 3
US-09-471-803A-6

Sequence 6, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471,803A
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnf
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brne
OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAACTGATTTCTCTGTAATCTCTTGTGTGCGAGATGTGAGTCATTACTTTT 60
DB 853 ATGCAACTGATTTCTCTGTAATCTCTTGTGTGCGAGATGTGAGTCATTACTTTT 912
QY 61 GCGCTCCGGGGGGGTTCCGTTCTTAATCCTTAAGCCCTTACGTGATCACAATTTGTGGC 120
DB 913 GCGCTCCGGGGGGGTTCCGTTCTTAATCCTTAAGCCCTTACGTGATCACAATTTGTGGC 972
QY 121 AAAATGGCGATGTGATGCGACGAGAGATCTTGGCCATTGTGACCGCATCAAGCTTTCG 180
DB 973 AAAATGGCGATGTGATGCGACGAGAGATCTTGGCCATTGTGACCGCATCAAGCTTTCG 1032
QY 181 AGCAATGCGATGATGATGAGATCTTAACCTTGTGTCATGTCGCTGTGAGTTACAGTG 240
DB 1033 AGCAATGCGATGATGATGAGATCTTAACCTTGTGTCATGTCGCTGTGAGTTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGACGACCTGTGTGAGCGTTGGCGGTGACCACTGCT 1152
QY 301 TTGTGTGACTGTGATCTTTTC 324
DB 1153 TTGTGTGACTGTGATCTTTTC 1176

RESULT 4

US-09-252-991A-13729/c
Sequence 13729, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 27.2807 Seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324
Sequence: 1 atgacaactgattctctctg.....ttgactgctgactctcttc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	4	US-09-471-803A-4
2	324	100.0	1271	4	US-09-471-803A-1
3	324	100.0	1271	4	US-09-471-803A-6
4	34.2	10.6	594	4	US-09-252-991A-13729
5	34.2	10.6	678	4	US-09-252-991A-13832
6	31.6	9.8	1404	4	US-09-328-352-1339
7	31.4	9.7	2387	4	US-08-961-527-56
8	30.8	9.5	5829	4	US-08-809-254A-4
9	30.6	9.4	7218	4	US-08-222-463-14
10	29.8	9.2	4635	2	US-08-404-531B-27
11	29.8	9.2	4635	3	US-08-476-900A-27
12	29.8	9.2	4635	3	US-08-488-546A-27
13	29.8	9.2	5110	2	US-08-404-531B-5
14	29.8	9.2	5110	2	US-08-404-531B-5
15	29.8	9.2	5110	3	US-08-476-900A-4
16	29.8	9.2	5110	3	US-08-476-900A-5
17	29.8	9.2	5110	3	US-08-488-546A-4
18	29.8	9.2	5110	3	US-08-488-546A-5
19	29.6	9.1	1308	2	US-08-404-531B-2
20	29.6	9.1	1308	2	US-08-404-531B-2
21	29.6	9.1	1308	2	US-08-476-900A-1
22	29.6	9.1	1308	3	US-08-476-900A-2
23	29.6	9.1	1308	3	US-08-488-546A-1
24	29.6	9.1	1308	3	US-08-488-546A-2
25	29.6	9.1	2454	2	US-08-404-531B-32
26	29.6	9.1	2454	3	US-08-476-900A-32
27	29.6	9.1	2454	3	US-08-488-546A-32

C	28	29.6	9.1	4931	3	US-08-726-320-2	Sequence 2, Appli
C	29	29.6	9.1	4931	3	US-09-208-716-2	Sequence 2, Appli
C	30	29.4	9.1	611	4	US-09-976-594-152	Sequence 152, App
C	31	29.4	9.1	12588	2	US-08-387-942C-1	Sequence 1, Appli
C	32	29.2	9.0	867	4	US-09-216-393B-340	Sequence 340, App
C	33	29.2	9.0	867	4	US-09-216-393B-342	Sequence 342, App
C	34	29.2	9.0	1397	4	US-09-216-393B-343	Sequence 343, App
C	35	29.2	9.0	1397	4	US-09-216-393B-345	Sequence 345, App
C	36	29.2	9.0	9828	4	US-08-961-527-41	Sequence 41, Appli
C	37	28.8	8.9	505	4	US-09-621-976-15639	Sequence 15639, A
C	38	28.4	8.8	459	2	US-08-387-942C-35	Sequence 35, Appli
C	39	28.4	8.8	7304	4	US-09-453-702B-174	Sequence 174, App
C	40	28.2	8.7	387	4	US-08-651-155B-74	Sequence 74, Appli
C	41	28.2	8.7	387	4	US-08-194-036B-74	Sequence 74, Appli
C	42	28.2	8.7	5099	4	US-09-610-040-5	Sequence 5, Appli
C	43	28.2	8.7	1230025	4	US-09-198-452A-1	Sequence 1, Appli
C	44	28	8.6	531	4	US-08-671-548C-7	Sequence 7, Appli
C	45	28	8.6	1758	4	US-09-205-258-173	Sequence 173, App

ALIGNMENTS

RESULT 1
US-09-471-803A-4
Sequence 4, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMANN

APPLICANT: EGGELING, LOTMAR

APPLICANT: PFEFFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

FILE REFERENCE: 21123/265496/MAS

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 4

LENGTH: 324

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(324)

OTHER INFORMATION: brne

OTHER INFORMATION: ATCC14752

US-09-471-803A-4

Query Match

Best Local Similarity

Matches 324; Conservative

0; Mismatches

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

CC The present invention relates to oligonucleotide clones originating in
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX Sequence 470 BP; 122 A; 131 C; 102 G; 106 T; 0 U; 9 Other;

SQ Query Match 9.9%; Score 32; DB 8; Length 470;

Best Local Similarity 51.9%; Pred. No. 2.2;

Matches 68; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY	181	ACCAATGCGATGATCTGAAGACTTAACCTTTGCTCTCATGCGCTTGCGATTACAGTG	240
DB	384	ATCAAGNACAGGAGNAGCATTAAGGAACTTTGAGNCCATCCACGTTGAGATGCCAG	325
QY	241	GTGGCGCATCTTTGGCGGCTCGACGACCTTTGTTGAGCGTTGGCGCTGCACCATCGTT	300
DB	324	GGTACACGATTTGCCACCCGCTGGGCAAGCTGTTCAACATTGSCAAGGCAACCAAGCT	265
QY	301	TTGTTGACT	311
DB	264	TGGGTGAGCT	254

Search completed: April 19, 2004, 11:42:06
Job time : 151.886 secs

CC The present invention relates to oligonucleotide clones originating in
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX Sequence 376 BP; 97 A; 105 C; 82 G; 92 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 376;
Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTCTAACCTTTGCTCATTTGCCGTTGCGATTACAGTGTGGCG 246
DB 375 GAGAGAGAGAGAGCATTAAGGAACTTTGAGACCATCCAGTTCAGATGCCAGGGTCAC 316
QY 247 CATCTTTGGCGGTCACGACCTTTGTAGCGTTGGCGCTGGCAACCATGTTTGT 306
DB 315 CAGTTTGCACCCGCTCTGGGCAACGTTTCACCATTTGGCAAGGACCAAGCCTTGGGTG 256
QY 307 GGAAT 311
DB 255 AGCCT 251

RESULT 14
ACL20868/c
ID ACL20868 standard; DNA; 392 BP.

XX ACL20868;

DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10859.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KM gene; ss.

XX *Hordeum vulgare*; var. (cul. Akashinriki).

XX WO2003057877-A1.

XX 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 20-DEC-2001; 2001JP-00403300.

PR 27-SEP-2002; 2002JP-00327515.

XX (UNNT-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA
XX sequences containing them for analysis and identification of barley
XX varieties and production of barley transformants with desired
XX characteristics.

XX Disclosure; SEQ ID XX; 284bp; Japanese.

CC The present invention relates to oligonucleotide clones originating in
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX Sequence 392 BP; 100 A; 105 C; 85 G; 102 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 392;
Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTCTAACCTTTGCTCATTTGCCGTTGCGATTACAGTGTGGCG 246
DB 391 GAGAGAGAGAGCATTAAGGAACTTTGAGACCATCCAGTTCAGATGCCAGGGTCAC 332
QY 247 CATCTTTGGCGGTCACGACCTTTGTAGCGTTGGCGCTGGCAACCATGTTTGT 306
DB 331 CAGTTTGCACCCGCTCTGGGCAACGTTTCACCATTTGGCAAGGACCAAGCCTTGGGTG 272
QY 307 GGAAT 311
DB 271 AGCCT 267

RESULT 15
ACL20918/c
ID ACL20918 standard; DNA; 470 BP.

XX ACL20918;

DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10909.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KM gene; ss.

XX *Hordeum vulgare*; var. (cul. Haruna Nijo).

XX WO2003057877-A1.

XX 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

XX 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 20-DEC-2001; 2001JP-00403300.

PR 27-SEP-2002; 2002JP-00327515.

XX (UNNT-) UNIV JAPAN OKAYAMA.

XX Sato K, Takeda K, Kohara Y;

XX WPI; 2003-587127/55.

XX Single nucleotide polymorphism sites in barley varieties and DNA
XX sequences containing them for analysis and identification of barley
XX varieties and production of barley transformants with desired
XX characteristics.

XX Disclosure; SEQ ID XX; 284bp; Japanese.

Query Match 10.0%; Score 32.4; DB 6; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCTACGTAATCAATTTGGGCAAAATGGCGATGTGATGCCAGAGAAATCCT 152
 DB 5382 GCACACACGTCATGTGCACATGTGTGAAAGAGGCTGGTGAAGACATCAATCCT 5323
 QY 153 TGGCATTGTGACCGCATCAAGTTTGGCAGCAATGGGATGATCTGAAGACTTAACCTT 212
 DB 5322 TGTCACTGGC--CAACAGAAATGTGCAAGATGGGCTGGTGTGTGTTCACACTG 5265
 QY 213 TGGTCTCA 220
 DB 5264 TAATCCCA 5257

RESULT 12
 ACF39600/c
 ID ACF39600 standard; cDNA; 6672 BP.
 XX AC ACF39600;
 XX DT 29-SEP-2003 (first entry)
 XX DE Human MHC class II transactivator encoding cDNA SEQ ID NO:3.
 XX
 XX Human; major histocompatibility complex class II transactivator;
 XX MHC class II transactivator; antisense modulation; immunosuppressive;
 XX antimicrobial; antidiabetic; antirheumatic; antitachytic; cytostatic;
 XX neurotropic; neuroprotective; immunostimulant; autoimmune disorder;
 XX MHC class II transactivator inhibitor; infection; transplant rejection;
 XX diabetes; rheumatoid arthritis; cancer; Alzheimer's disease;
 XX multiple sclerosis; severe combined immunodeficiency disease; gene; ss.
 XX OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX FT 139..3531
 XX FT CDS /*tag= a
 XX /product= "MHC class II transactivator"
 XX
 XX MO2003050247-A2.
 XX PD 19-JUN-2003.
 XX PF 04-DEC-2002; 2002MO-US038616.
 XX PR 05-DEC-2001; 2001US-00006366.
 XX PA (ISIS-) ISIS PHARM INC.
 XX PI Bennett FC, Dobie KM;
 XX DR WPI; 2003-577294/54.
 XX DR P-PSDB; ABR82108.
 XX PT New antisense oligonucleotides for modulating MHC class II transactivator
 XX gene expression, particularly useful for treating autoimmune disorders,
 XX PT such as transplant rejection, Alzheimer's disease, or multiple sclerosis,
 XX PT or infection.
 XX
 XX Example 13; Page 93-99; 129pp; English.
 XX
 XX The present invention describes a compound (I) that is 8-50 nucleobases
 XX in length: (a) targets a nucleic acid molecule encoding major
 XX histocompatibility complex (MHC) class II transactivator, and
 XX specifically hybridizes with the nucleic acid encoding the MHC class II
 XX transactivator, and inhibits the expression of MHC class II
 XX transactivator; or (b) specifically hybridizes with at least an 8-
 XX nucleobase portion of an active site on a nucleic acid molecule encoding
 XX MHC class II transactivator. (I) has immunosuppressive, antimicrobial,
 XX antidiabetic, antirheumatic, antitachytic, cytostatic, neurotropic,

CC neuroprotective and immunostimulant activities, and can be used as an MHC
 CC class II transactivator inhibitor. The MHC class II transactivator
 CC antisense oligonucleotides can be used for treating an animal having a
 CC disease or condition associated with MHC class II transactivator, e.g.,
 CC autoimmune disorder or infection. The antisense oligonucleotides can be
 CC used for inhibiting the expression of MHC class II transactivator in
 CC cells or tissues. In particular, these diseases include transplant
 CC rejection, diabetes, rheumatoid arthritis, cancer, Alzheimer's disease,
 CC multiple sclerosis, or severe combined immunodeficiency disease. The
 CC antisense compounds are useful for diagnostics, prophylaxis, or as
 CC research reagents or kits. The present sequence encodes a human MHC class
 CC II transactivator, which is used in an example from the present invention
 XX

Sequence 6672 BP; 1392 A; 1954 C; 1868 G; 1458 T; 0 U; 0 Other;

Query Match 10.0%; Score 32.4; DB 8; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCTACGTAATCAATTTGGGCAAAATGGCGATGTGATGCCAGAGAAATCCT 152
 DB 5382 GCACACACGTCATGTGCACATGTGTGAAAGAGGCTGGTGAAGACATCAATCCT 5323
 QY 153 TGGCATTGTGACCGCATCAAGTTTGGCAGCAATGGGATGATCTGAAGACTTAACCTT 212
 DB 5322 TGTCACTGGC--CAACAGAAATGTGCAAGATGGGCTGGTGTGTGTTCACACTG 5265
 QY 213 TGGTCTCA 220
 DB 5264 TAATCCCA 5257

RESULT 13
 ACL20925/c
 ID ACL20925 standard; DNA; 376 BP.
 XX AC ACL20925;
 XX DT 27-OCT-2003 (revised)
 XX DT 17-OCT-2003 (first entry)
 XX DE DNA clone originating in barley containing SNP encoding sequence #10916.
 XX
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 XX gene; ss.
 XX KW Hordeum vulgare; var. (cul. Akashinrkt).
 XX OS
 XX PN WO2003057877-A1.
 XX PD 17-JUL-2003.
 XX PF 16-DEC-2002; 2002MO-IB005403.
 XX PR 20-DEC-2001; 2001JP-00387059.
 XX PR 20-DEC-2001; 2001JP-00387131.
 XX PR 20-DEC-2001; 2001JP-00403299.
 XX PR 27-SEP-2002; 2002JP-00387515.
 XX
 XX (UNN-) UNIV JAPAN OKAYAMA.
 XX PI Sato K, Takeda K, Kohara Y;
 XX DR WPI; 2003-587127/55.
 XX PT Single nucleotide polymorphism sites in barley varieties and DNA
 XX PT sequences containing them for analysis and identification of barley
 XX PT varieties and production of barley transformants with desired
 XX characteristics.
 XX PS Disclosure; SEQ ID XX; 284dp; Japanese.
 XX

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB57237-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO.int/publ/publ/published_pct_sequences

XX Sequence 3144 BP; 1056 A; 676 C; 680 G; 732 T; 0 U; 0 Other;

Query Match 10.1%; Score 32.8; DB 4; Length 3144;
Best Local Similarity 52.1%; Pred. No. 2.8; Mismatches 0; Gaps 0;
Matches 73; Conservative 0; Indels 67; Indels 0; Gaps 0;

169 TCACGTTTGGCAGCATGCGATGATCTGAGACTCTACCTTGGTCTGATGCGCT 228
1512 TCCTAATTTCTATATATATTTAGTTAAATGATGATTTATGTTTACCTGATATCTGTT 1453

229 GCGATTACAGTGTGGCGCATCTTCTGGCGTCGACGCACTGTTGAGCGTGGCGCT 288

1452 TCGTTTGGGGTGTGGTGTCTTCTTTCGTCGCGGAGCAATGCTTGTGCGTGGCA 1393

289 GGCACCATCGTTTTTGTGG 308

1392 CACAGTCTTATGTTGTGG 1373

RESULT 9
AAS59515/C
ID AAS59515 standard; DNA; 66788 BP.

XX AAS59515;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #10.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhactia A,
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris.

XX Claim 1; SEQ ID NO 10; 1069bp; English.

CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by *P. acnes*. The disorders include

CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of *P. acnes* in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for *P. acnes* proteins. These antibodies can be
CC used to downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU1313-4209 and AAU67473. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO.int/publ/publ/published_pct_sequences

XX Sequence 66788 BP; 13534 A; 21369 C; 19347 G; 12530 T; 0 U; 8 Other;

Query Match 10.1%; Score 32.6; DB 4; Length 66788;
Best Local Similarity 52.6%; Pred. No. 13; Mismatches 64; Indels 0; Gaps 0;
Matches 71; Conservative 0; Indels 64; Indels 0; Gaps 0;

154 GCCATTTGACCGCATCAACGTTTGGCAGCATGATCTGAAAGACTTAACTTT 213

4085 GTCAATTCGATCGGGGAGCAACCGCCGACGTTCTTGACCGGCTGCTACGTCCT 40026

214 GGTTCATTTGCGCGTTCGATTCAGTGTGGCGCATCTTCTGGCGGTTCACGCACTTG 273

40025 GGGCGTTTCGACCGTCGATGAGTGGGCTCGGATTCGATGGCGCTCAAGATT 39966

274 TTGAGCGTTGGCGCT 288

39965 CTCAGGTTCAAGCT 39951

RESULT 10
ACF64444/C

ID ACF64444 standard; DNA; 66788 BP.

XX ACF64444;

XX 17-OCT-2003 (first entry)

XX Propionibacterium acnes DNA contig sequence #10.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine; ds.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAM, Persing DH, Bhactia A, Maisonneuve JL,
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX Barch B, Valliave-Douglass J;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a *P. acnes* protein.

XX Claim 1; SEQ ID NO 10; 1481bp; English.

CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;

Query Match 100.0%; Score 324; DB 5; Length 349980;

Best Local Similarity 100.0%; Pred. No. 2.1e-99;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACACGATGATTCCTGCTATTCCTGCTGTCGACGATGATGACGATTAATCTTT 60
DB 277581 ATGACACGATGATTCCTGCTATTCCTGCTGTCGACGATGATGACGATTAATCTTT 277640
QY 61 GCGCTCGGCGGCTTCGCTTCTTAATCCTTAAGCCCCCTAAGTATCACAATTTGTGGC 120
DB 277641 GCGCTCGGCGGCTTCGCTTCTTAATCCTTAAGCCCCCTAAGTATCACAATTTGTGGC 277700
QY 121 AAAATGCGCATGTGATGATGACGACGACGAAATCCTTGAACCGGATCAACGTTTGGC 180
DB 277701 AAAATGCGCATGTGATGATGACGACGACGAAATCCTTGAACCGGATCAACGTTTGGC 277760
QY 181 AGCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 277761 AGCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277820
QY 241 GTGGCGCATGTTTGTGGCGGTGACGACGACGATGTTGAGCGTTGGCGTGGCAACATCGTT 300
DB 277821 GTGGCGCATGTTTGTGGCGGTGACGACGACGATGTTGAGCGTTGGCGTGGCAACATCGTT 277880
QY 301 TTGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 324
DB 277881 TTGTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277904

```

RESULT 7
 ADA71938/c
 ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.

XX Oryza sativa.

XX WO200300898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899BP; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 10.7%; Score 34.8; DB 7; Length 2000;

Best Local Similarity 12.9%; Pred. No. 0.47;

Matches 30; Conservative 105; Mismatches 97; Indels 0; Gaps 0;

```

QY 4 ACAACTGATTTCTCTGATTCCTGCTGTCGACGATGATGACGATTAATCTTTGCG 63
DB 901 AYAARRRRATYKMSWRMYWTMTKMAWTMTCMCKMYATGATMMWRYYTYCY 842
QY 64 CTCGGGGGCTTCGCTTCTTAATCCTTAAGCCCTAGCGATGATGATGATGATGATGATG 123
DB 841 AATCAKCKYKMAATKMTTMAACMPATSWRMANAKKRYKRAVMMRWKQWRAW 782
QY 124 ATGGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 183
DB 781 ARWMSRRMYKMYATYKMAATMMWMSRWKSYRMSGMRMSAMRYCSRMKC 722
QY 184 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 235
DB 721 AATKYASARWTKAKRSYRRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 670

```

RESULT 8

ABL27130/c
 ID ABL27130 standard; DNA; 3144 BP.

XX ABL27130;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 32863.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

XX Claim 1; SEQ ID NO 32863; 21BP + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

Oy		301	TTCGTTGGACTGTGAATCTTTC	324
Dd		1153	TTCGTTGGACTGTGAATCTTTC	1176

```
RESULT 5
AAH21112
ID    AAH21112 standard; DNA; 1271 BP.
...
```

DT 05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnE.

KM l-amino acid production; brnF; brnE; branched-chain amino acid;
KM coryneform bacterium; leucine; isoleucine; valine; medicine;
KM animal nutrition; ds.

Corynebacterium glutamicum.

FH	Key	Location/Qualifiers
7E	07C	101 855

[illegible]

FT	CDS	853. .1179
CE		4433- b

PN EP1096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

PR 27-OCT-1999; 99DE-01051708.

PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

PI Kennerknecht N, Eggeling L, Sahm H, Pfefferle W;

DR WPI; 2001-391595/42.

PT New export genes from coryneform bacteria, useful for increasing fermentative production of branched-chain amino acids.

PS Claim 4 (1); Page 17-18; 23pp; German.

This invention describes a novel isolated polynucleotide (1) containing at least one sequence that (i) is 70% identical with a sequence that encodes at least one of 251 amino acid (aa), (3) or 108 aa (5) polypeptides; (ii) encodes a polypeptide at least 70% identical with (3) or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least 15 consecutive bases from (i)-(iii). The invention also describes (a) protein derived from sequences of 753 bp (2) or 324 bp (4); (b) Corynebacterium microorganisms, especially Corynebacterium, transformed with one or more (1), where these are replicative DNA; (c) production of branched-chain l-a-a by fermentation of corynebacter bacteria in which the bnf and/or bnf genes (or equivalent sequences) are amplified, especially overexpressed; and (d) method for isolating the bnf and/or bnf genes. (1) is used for transformation of corynebacter bacteria being used for fermentative production of branched-chain amino acids, specifically leucine, isoleucine and valine, which are useful in medicine and animal nutrition. (1) can also be used as source of primers and probes for isolation of related sequences. Transformation with (1) increases yield of branched-chain amino acids. This sequence encodes the Corynebacterium glutamicum ATCC 13032 bnf and bnf proteins described in the method of the invention

SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match	100.0%	Score 324;	DB 4;	Length 1271;
Best Local Similarity	100.0%	Pred. No. 1.6e-100;		
Matches 324;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	ATGCAACTGATATTTCTCCTGATATCTCTTGTTGTGCGAGATATGTGCAGTCACTACTTTT	60
Db	ATGCAACTGATATTTCTCCTGATATCTCTTGTTGTGCGAGATATGTGCAGTCACTACTACTTTT	853
QY	GCCTCCGGCGGGTTCCTGTTCTTAATCCTTAAAGCCCTTACGTGAATATACAAATTTGTGGC	120
Db	GCCTCCGGCGGGTTCCTGTTCTTAATCCTTAAAGCCCTTACGTGAATATACAAATTTGTGGC	913
QY	AAATGCGAGATGTGATAGCCAGCAGAGATCCTTGCATTTTGAACCGCATCAACGTTTGGC	180
Db	AAATGCGAGATGTGATAGCCAGCAGAGATCCTTGCATTTTGAACCGCATCAACGTTTGGC	973
QY	AGCAATGCGAATATCTGAAAGACTCTTAACCTTGTGATTCATTTGCCTGTGCGATTAACAATG	240
Db	AGCAATGCGAATATCTGAAAGACTCTTAACCTTGTGATTCATTTGCCTGTGCGATTAACAATG	1033
QY	GTGGCGCATCTCTCTTGGCGGTGACGACGACCTTTGTGAAGCGTTGGCGGTGGACCAATCGTT	300
Db	GTGGCGCATCTCTCTTGGCGGTGACGACGACCTTTGTGAAGCGTTGGCGGTGGACCAATCGTT	1093
QY	TTTGTGGACTGTGAATCTTTTC	324
Db	TTTGTGGACTGTGAATCTTTTC	1153

RESULT 6
AAH64966
ID AAH64966 standard; DNA; 349980 BP

DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 1.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KN organic acid synthesis; ds.

Corynebacterium glutamicum

PN EP1108790-A2.

PD 20-JUN-2001

PF 18-DEC-2000; 2000EP-001276BB
XX

PR	16-DEC-1999;	99JP-00377484
PR	07-APR-2000;	2000JP-00159162

03-AUG-2000; 2000JP-00280988
PR
XX

PA (KICM) KIOMH HANAO KOGIO NA
XX

PI Nakagawa S, Mizoguchi H, Arai T, Tateishi N, Senoh A, Ikeda PI

WPI: 2001-376931/40.

Novel polynucleotides derived from

PT. expression profile or pattern

PS Claim 7; SEQ ID NO 1; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein

CC sequences from the *Corynebacterium* bacterium *Corynebacterium glutamicum*, these
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of *Corynebacterium* bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC *Corynebacterium* bacterium, and identifying a homologue of a gene derived from

XX		PR	13-JUN-2001; 2001DE-01028510.
XX	PA	(DEGS) DEGUSA AG.	
XX	P1	Farwick M, Moekkel B, Pfeifferle W, Bathe B, Huthmacher K;	
XX	DR	WPI; 2003-279970/28.	
XX			
PT		New nucleic acid array useful for monitoring mRNA expression of	
PT	Corynebacterium glutamicum during fermentation, comprising nucleic acid		
FT	from Corynebacterium glutamicum.		
XX			
PS	Claim 1; Page 647; 709pp; German.		
XX			
CC	This invention describes a novel nucleic acid array involving		
CC	Corynebacterium glutamicum polynucleotides. The arrays are used to		
CC	analyse C. glutamicum, particularly for monitoring a fermentation process		
CC	to determine expression levels of C. glutamicum cellular mRNA. Such		
CC	monitoring particularly differentiates between expression levels of		
CC	different strains of C. glutamicum and allows the adjustment of different		
CC	culture and fermentation conditions. ACAA0010-ACAQ2188 represent C.		
CC	glutamicum derived polynucleotides described in the disclosure of the		
CC	Invention		
XX			
SQ	Sequence 324 BP; 60 A; 79 G; 80 T; 105 U; 0 Other;		
	Query Match 100.0%; Score 324; DB 7; Length 324;		
	Best Local Similarity 100.0%; Pred. No. 8.8e-101;		
	Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
OY	1 ATGACAACGTATTTCTTCGTGAATCTCCTTGGTGCGAGATGTGCAGTCAATTACTTTT	60	
Dd	1 ATGACAACGTATTTCTTCGTGAATCTCCTTGGTGCGAGATGTGCAGTCAATTACTTTT	60	
OY	61 GGCGTCCGGGGGGGTTCCGTTCTTAATCCTTAGCCCCACGAAATCAATTGAGGGC	120	
Dd	61 GGCGTCCGGGGGGGTTCCGTTCTTAATCCTTAGCCCCACGAAATCAATTGAGGGC	120	
OY	121 AAAATGGCGATGTGATGTCGACAGAAGAATCCTTGCATTTTGACCGCATCAAGTTTTGC	180	
Dd	121 AAAATGGCGATGTGATGTCGACAGAAGAATCCTTGCATTTTGACCGCATCAAGTTTTGC	180	
OY	181 AGCATGCGATGATGATCTGAAGCTTAACCTTGGTCTCATTCGCCGTGCCATTACAGTG	240	
Dd	181 AGCATGCGATGATGATCTGAAGCTTAACCTTGGTCTCATTCGCCGTGCCATTACAGTG	240	
OY	241 GTGGGCGCATCTTCTTGCGSGTGAAGCACCTTGTTGAGCGGTGGCGTGGCACTACGTT	300	
Dd	241 GTGGGCGCATCTTCTTGCGSGTGAAGCACCTTGTTGAGCGGTGGCGTGGCACTACGTT	300	
OY	301 TTGTGTTGACTGTGATACTTTTTC	324	
Dd	301 TTGTGTTGACTGTGATACTTTTTC	324	
RESULT 4			
AHH21109			
ID	AHH21109 standard; DNA; 1271 BP.		
XX			
AC	AHH21109;		
DT	05-SEP-2001 (first entry)		
XX			
DE	C. glutamicum DNA encoding brnF and brnE.		
XX			
KX	L-amino acid production; brnF; brnE; branched-chain amino acid;		
KM	coryneform bacterium; leucine; isoleucine; valine; medicine;		
XX	animal nutrition; ds.		
OS	Corynebacterium glutamicum.		
XX			
Key	Location/Qualifiers		
PH	CDS		
FT	101..856		

FT	/*tag= a	
FT	/product= "brnF"	
FT	653..1179	
FT	CDS	
FT	/*tag= b	
FT	/product= "brnE"	
FN	EP1096010-A1.	
PD		
PD	02-MAY-2001.	
PF	11-OCT-2000; 2000EP-00122057.	
PR	27-OCT-1999; 99DE-01051708.	
PA	(DECS) DEGUSA AG.	
PA	(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.	
PI	Kemmerknecht N, Eggeling L, Sahm H, Pfeffeler W;	
DR	WPI: 2001-391595/42.	
DR	P-PSDB; AAB86247, AAB86248.	
XX		
PT	New export genes from coryneform bacteria, useful for increasing	
PT	fermentative production of branched-chain amino acids.	
PS		
PS	Claim 4 (1); Page 13; 23pp; German.	
XX		
XX	This invention describes a novel isolated polynucleotide (1) containing	
CC	at least one sequence that (1) is 70% identical with a sequence that	
CC	encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)	
CC	polypeptides; (11) encodes a polypeptide at least 70% identical with (3)	
CC	or (5); (111) is the complement of (1) or (11), or (1v) contains at least	
CC	15 consecutive bases from (1)-(111). The invention also describes (a)	
CC	protein derived from sequences of 753 bp (2) or 324 bp (4); (b)	
CC	coryneform microorganisms, especially Corynebacterium, transformed with	
CC	one or more (1), where these are replicative DNA; (c) production of the	
CC	branched-chain L-aa by fermentation of coryneform bacteria in which the	
CC	brnE and/or brnF genes (or equivalent sequences) are amplified,	
CC	especially overexpressed; and (d) method for isolating the brnE and/or	
CC	brnF genes. (1) is used for transformation of coryneform bacteria being	
CC	specifically fermentative production of branched-chain amino acids;	
CC	used for fermentative production of branched-chain amino acids;	
CC	and animal nutrition. (1) can also be used as source of primers and	
CC	probes for isolation of related sequences. Transformation with (1)	
CC	increases yield of branched-chain amino acids. This sequence encodes the	
CC	Corynebacterium glutamicum ATCC 14752 brnF and brnE proteins described in	
CC	the method of the invention	
XX		
XX		
SQ	Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;	
	Query Match	100.0%; Score 324; DB 4; Length 1271;
	Best Local Similarity	100.0%; Pred. No. 1.6e-100;
	Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGACACTGATTTCTCCTGATATCTCCTGATCTCCTGTTGTCGACAGTATGAGATCAATCTTT	60
DB	853 ATGCACTAGATTTCTCTCTGATTTCTCTGTTGTTGTCGACAGTATGAGATCAATCTTT	912
QY	61 GCGCTCGGGGGCGGCTTCGTTCTTATCCTTAAGCCCTACGTCGATCAATCAATTTGTGGC	120
DB	913 GCGCTCGGGGGCGGCTTCGTTCTTATCCTTAAGCCCTACGTCGATCAATTTGTGGC	972
QY	121 AAATGCGGAGTGGATATGCCAGCGAGATCCTTGGCATTTTGACCGCATCAAGTTTGGC	180
DB	973 AAATGCGGAGTGGATATGCCAGCGAGATCCTTGGCATTTTGACCGCATCAAGTTTGGC	1032
QY	181 AGCAATGCGATGATCTGAAGACTCTAACCTTTGTCTCATTTGGCGATTACAGT	240
DB	1093 AGCAATGCGATGATCTGAAGACTCTAACCTTTGTCTCATTTGGCGATTACAGT	10922
QY	241 GTGGCGCATCTTTCTTGGCGGTCGACGCACTTTGTTAGCGTTGGCGCTGGACCATGTT	300
DB	1093 GTGGCGCATCTTTCTTGGCGGTCGACGCACTTTGTTAGCGTTGGCGCTGGACCATGTT	1152

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 150.886 Seconds

(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Sequence: 324 1 atgacaactgattctctctg.....cttgactgtgactctcttc 324

Scoring table: IDENTITY NUC

Gapop 10_0, Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018s:*
6: geneseqn20028s:*
7: geneseqn20038s:*
8: geneseqn20038s:*
9: geneseqn20038s:*
10: geneseqn20048s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	AAH21111	AAH21111 C. glutam
2	324	100.0	324	AAH65255	AAH65255 C. glutam
3	324	100.0	324	ACA01970	ACA01970 C. glutam
4	324	100.0	1271	AAH21109	AAH21109 C. glutam
5	324	100.0	1271	AAH21112	AAH21112 C. glutam
6	324	100.0	349980	AAH64966	AAH64966 C. glutam
7	34.8	10.7	2000	ADA71938	ADA71938 Rice gene
8	32.8	10.1	3144	ABL27130	ABL27130 Drosophila
9	32.6	10.1	66788	AAH59515	AAH59515 Propionib
10	32.6	10.1	66788	ACF64444	ACF64444 Propionib
11	32.4	10.0	6672	ABH68121	ABH68121 Ovary can
12	32.4	10.0	6672	ACF39600	ACF39600 Human NEC
13	32.4	9.9	376	ACT20925	ACT20925 DNA clone
14	32.2	9.9	392	ACL20868	ACL20868 DNA clone
15	32	9.9	470	ACL20918	ACL20918 DNA clone
16	32	9.9	476	AAH20933	AAH20933 DNA clone
17	32	9.9	820	AAH05013	AAH05013 Human CDN
18	32	9.9	1808	AAH18459	AAH18459 Human CDN
19	31.8	9.8	2091	ADB06513	ADB06513 Altiococ
20	31.8	9.8	2091	ADB06509	ADB06509 Altiococ
21	31.8	9.8	2091	ADB06511	ADB06511 Altiococ
22	31.8	9.8	110000	ADB12064_00	ADB12064 Altiococ
23	31.8	9.8	110000	ADB12064_01	ADB12064 Altiococ

ALIGNMENTS

24	31.6	9.8	1404	8	ADA30052	ADA30052 DNA encod
25	31.4	9.7	2387	2	AAV52189	AAV52189 Streptoco
26	31.4	9.7	110000	7	ABH56454_07	Continuation (8 of
27	31	9.6	4590	5	AAH24065	AAH24065 Yeast AOD
28	31	9.6	6415	2	AAH13212	AAH13212 Enterococ
29	31	9.6	6415	6	ABH99007	ABH99007 Enterococ
30	30.8	9.5	5949	2	AAT30053	AAT30053 S. pneumo
31	30.8	9.5	265118	5	AAH41227	AAH41227 Pyrococu
32	30.6	9.4	528	8	ACL20863	ACL20863 DNA clone
33	30.6	9.4	789	7	ACA50939	ACA50939 Prokaryot
34	30.4	9.4	386	8	ACL20938	ACL20938 DNA clone
35	30.4	9.4	4739	6	ABH03452	ABH03452 Negative-
36	30.2	9.3	373	8	ACL20861	ACL20861 DNA clone
37	30.2	9.3	448	8	ACL20929	ACL20929 DNA clone
38	30.2	9.3	505	8	ACL20950	ACL20950 DNA clone
39	30.2	9.3	515	8	ACL20954	ACL20954 DNA clone
40	30.2	9.3	528	8	ACL20920	ACL20920 DNA clone
41	30.2	9.3	528	8	ACL20912	ACL20912 DNA clone
42	30.2	9.3	533	8	ACL20906	ACL20906 DNA clone
43	30.2	9.3	537	8	ACL20947	ACL20947 DNA clone
44	30.2	9.3	538	8	ACL20907	ACL20907 DNA clone
45	30.2	9.3	540	8	ACL20924	ACL20924 DNA clone

RESULT 1
AAH21111
ID AAH21111 standard; DNA; 324 BP.

AAH21111;
05-SEP-2001 (first entry)
C. glutamicum brne DNA.

L-amino acid production; brnf; brng; branched-chain amino acid;
coryneform bacterium; leucine; isoleucine; valine; medicine;
animal nutrition; ds.

Corynebacterium glutamicum.

EPI096010-A1.

02-MAY-2001.

11-OCT-2000; 2000EP-0012057.

27-OCT-1999; 99DE-01051708.

(DEGS) DEGUSA AG.

(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

Kennerknecht N, Eggeling L, Sahm H, Pfefferle W;

WPI; 2001-391595/42.

P-PSDB; AAB86248.

New export genes from coryneform bacteria, useful for increasing
fermentative production of branched-chain amino acids.

Claim 5; Page 16; 23pp; German.

This invention describes a novel isolated polynucleotide (I) containing
at least one sequence that (i) is 70% identical with a sequence that
encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least
15 consecutive bases from (1)-(iii). The invention also describes (a)
protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
coryneform microorganisms, especially Corynebacterium, transformed with
one or more (I), where these are replicative DNA; (c) production of

/product="putative mannosyltransferase A"
/protein_id="AA064208.1"
/db_xref="GI:28974207"
/translation="IHDLPITSPQPPPISSGCHLQWETIRKIDPVLATSOETKN
DLKRYGFKTKIPLGSNFTLKRIRKNTFLAVGTIERKQAKAVITFLQON
EGIEDLIVGKRGVDEDEKPEFLCKKRNIRYLFKVSDFLCSLYSTETALIASYD
EGYGLPMVVASYGACIIRADIPVFRDILGENATYKODISLYETIKYKSKTKRQI
GITWEYARIELIKLIEDL"
711..1856
/gene="mteB"
711..1856
/gene="mteB"
/note="Cf0002"
/codon_start=1
/transl_table=11
/product="putative mannosyltransferase B"
/protein_id="AA064209.1"
/db_xref="GI:28974208"
/translation="WINTLIDTPPLASNTGGRVYTYEIAKLNKNEPRTYDGFLL
DELIVDLMOQKRPISKHLINIAKILDKPTIKVYKMLIKNSLSEVSINID
YQPNPIPELAKAVYVTHDLAFMSYPEFKERIDYEQNFIPNLKTHPTVS
NAIKHGIKILKANSADKISVYNGVDENIFKPKQKOTINILNDRLNPFILFGSI
EPRKNTLTLQAVNELNNIDLVIVAGKEMSEIHLONNEHIFKGFPPDLA
TLVSATIFVYPSIGFGLPELAACGAPILSDIEVEREIVGNVAFSPPLNAKE
LAERKLNINSKLSIMKNGILSKSKYIMKCIQEHKLPILKL"
2182..5010
/gene="sapA"
2182..5010
/gene="sapA"
/note="Cf0003; similar to sap"
/codon_start=1
/transl_table=11
/product="surface array protein"
/protein_id="AA064210.1"
/db_xref="GI:28974209"
/translation="MLNKTDVSMLYITIMGASBGDNKYMLDYANNNTLGVSLANI
MLDPSGAKFPGDILAGNEKDFVTKYISALGNTSDVDGINWTAITGGSEFTSK
GNIVSVASLKGDLIGAMINSMVNGSASEKAIPEAKAASDFADATIGDISE
GTTSKLISEINSASDLIKVSEILDALSKSELPNGSVDLTDEGDNILKGDIDDTNGT
TYNGNINSLFSPDKIDGAGSDTINAOYLANNASDAKSEDLVQALAGVYVE
SVNIVDLPNDTALTNLSGSTMANALDVLSKLAGLTTDISGKUNIVNGSVDLT
AANKSVVAVAGVAVLTDTKAAINNVAKAGDVTLDLEPASISVDANGKSVTLT
KATVENLISYNAATNLITTAGDKNLTVLENTLIDSSAAAIDPFKATLLENFNV
KIDGDSKIAHISASSETLNPFSGEATFFALVTOVYANITANASNNFLIKKE
DGHDRDAALANDSFTFEVFKNGKELVFPAGLDVTSIDAGFNGKATVINGDNG
DATAGLVSGAGSDPTVLESNKLSAGSVLDGAGSDTILKSSALDABATIGMIRI
EVOVDKULANTINISTSAFTTIYAGDETSPTKLTINQKIDLAATTLAASKITM
DNEKSGDNDISINLVAKTADNNVTVNGAKOKVLTEIGISVNTISVAKSTIAN
SIDKFTDVKSIIINGDAGITLKGFTAVYKNSMIDASLAKGLTPASALITPANG
ATIKGSGADSITVKGGLANLTLGEGSDTTLKKGSTIDYITINIPSDKIVLDGT
DFAKAIEKLTANTTGTFSYVNOAASGGGNTPIVYKYPHYNDYIVVDKAGATF
TKATDVYIKLAGHELTGONTIIDSQ"
complement(5749..9069)
/gene="sap5"
complement(5749..9069)
/gene="sap5"
/note="Cf0004; similar to sap"
/codon_start=1
/transl_table=11
/product="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTDVSMLYITIMGASBGDNKYMLDYANNNTLGVSLANI
MLDPSGAKFPGDILAGNEKDFVTKYISALGNTSDVDGINWTAITGGSEFTSK
GNIVSVASLKGDLIGAMINSMVNGSASEKAIPEAKAASDFADATIGDISE
GTTSKLISEINSASDLIKVSEILDALSKSELPNGSVDLTDEGDNILKGDIDDTNGT
TYNGNINSLFSPDKIDGAGSDTINAOYLANNASDAKSEDLVQALAGVYVE
SVNIVDLPNDTALTNLSGSTMANALDVLSKLAGLTTDISGKUNIVNGSVDLT
AANKSVVAVAGVAVLTDTKAAINNVAKAGDVTLDLEPASISVDANGKSVTLT
KATVENLISYNAATNLITTAGDKNLTVLENTLIDSSAAAIDPFKATLLENFNV
KIDGDSKIAHISASSETLNPFSGEATFFALVTOVYANITANASNNFLIKKE
DGHDRDAALANDSFTFEVFKNGKELVFPAGLDVTSIDAGFNGKATVINGDNG
DATAGLVSGAGSDPTVLESNKLSAGSVLDGAGSDTILKSSALDABATIGMIRI
EVOVDKULANTINISTSAFTTIYAGDETSPTKLTINQKIDLAATTLAASKITM
DNEKSGDNDISINLVAKTADNNVTVNGAKOKVLTEIGISVNTISVAKSTIAN
SIDKFTDVKSIIINGDAGITLKGFTAVYKNSMIDASLAKGLTPASALITPANG
ATIKGSGADSITVKGGLANLTLGEGSDTTLKKGSTIDYITINIPSDKIVLDGT
DFAKAIEKLTANTTGTFSYVNOAASGGGNTPIVYKYPHYNDYIVVDKAGATF
TKATDVYIKLAGHELTGONTIIDSQ"
complement(5749..9069)
/gene="sap5"
complement(5749..9069)
/gene="sap5"
/note="Cf0004; similar to sap"
/codon_start=1
/transl_table=11
/product="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTDVSMLYITIMGASBGDNKYMLDYANNNTLGVSLANI
MLDPSGAKFPGDILAGNEKDFVTKYISALGNTSDVDGINWTAITGGSEFTSK
GNIVSVASLKGDLIGAMINSMVNGSASEKAIPEAKAASDFADATIGDISE
GTTSKLISEINSASDLIKVSEILDALSKSELPNGSVDLTDEGDNILKGDIDDTNGT
TYNGNINSLFSPDKIDGAGSDTINAOYLANNASDAKSEDLVQALAGVYVE
SVNIVDLPNDTALTNLSGSTMANALDVLSKLAGLTTDISGKUNIVNGSVDLT
AANKSVVAVAGVAVLTDTKAAINNVAKAGDVTLDLEPASISVDANGKSVTLT
KATVENLISYNAATNLITTAGDKNLTVLENTLIDSSAAAIDPFKATLLENFNV
KIDGDSKIAHISASSETLNPFSGEATFFALVTOVYANITANASNNFLIKKE
DGHDRDAALANDSFTFEVFKNGKELVFPAGLDVTSIDAGFNGKATVINGDNG
DATAGLVSGAGSDPTVLESNKLSAGSVLDGAGSDTILKSSALDABATIGMIRI
EVOVDKULANTINISTSAFTTIYAGDETSPTKLTINQKIDLAATTLAASKITM
DNEKSGDNDISINLVAKTADNNVTVNGAKOKVLTEIGISVNTISVAKSTIAN
SIDKFTDVKSIIINGDAGITLKGFTAVYKNSMIDASLAKGLTPASALITPANG
ATIKGSGADSITVKGGLANLTLGEGSDTTLKKGSTIDYITINIPSDKIVLDGT
DFAKAIEKLTANTTGTFSYVNOAASGGGNTPIVYKYPHYNDYIVVDKAGATF
TKATDVYIKLAGHELTGONTIIDSQ"

Query Match 11.8%; Score 38.2; DB 1; Length 65865;
Best Local Similarity 47.0%; Pred. No. 2.4;
Matches 118; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 CAACGATTTTCCTGCTTCTTCTTCTGTCGCAAGTATGCACTATTACTTTGGCC 64
DB 45075 CAACCTCCGTTGCGCTATTTTATTTTGTGCTGCACTTGGAGATTTGAGGCG 45016
65 TCCGGGCGGTCCTGCTTCTTAACTTAAGCCCTACGTAATCAAAATTTGGGCAAAA 124
DB 45015 TTTTACGTCCTCCGCTTGTTTTAAATCAACACTAGTAGAGCGTTGCACTGGTATG 44956
QY 125 TGGCGATGATGACGACGAGAACTCTGCACTTTGACCGCATCAACGTTGGACGA 184
DB 44955 TGATGATTCGCTTTTACAGTTAAATTTAGCGTGTGCGGAATCGGCGCTAAGGCGAG 44896
QY 185 ATGCGATGATCTGAAGCTTAACCTTGGCTCATTTGCCGTTGCGATTAAGTGATG 244
DB 44895 AATTTTACCGCGCAACATCGACTTTCCTGTTCCGCTGCGGTGAAGCAGTATCG 44836
QY 245 CGCATCTTCTT 255
DB 44835 TTCCTGTTGTT 44825
Search completed: April 19, 2004, 16:30:07
Job time : 1575.24 secs
NLSVTEGASFIATATSLAKNLNLSLNGVSADINSVNGCTAIIASLEANNVSGEFL
GTTAKGDVDNENIENVALTIGAITSGTGNASVYISATNNTLGAWSAQGLTLNA
GNTLGNITLAKGDIYSVDLGVLTGINSIDANNKYSITSENETYVSESKRVVEIT
AAGGDTLNQOVGGAADADALITIGGDTOTITASGDLGGTLTLTLEATLSSLD
ISGVKLSATATDIDKNSVENKNIYVNDIOSDAEITTAATTAITLSDLDGG
ANTYVAPDAAVAATTTIDSGSAGTGTSGTTHHAAOTATLTIKGSGNDPTTG
IANDLVTVGAGDVNFVPAKIVTANTEHATITDPSGDSITKPASTAYAKHSY
TLVSDLSLTANNIDIVFA"
complement(19374..10762)
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein cf0005"
/protein_id="AA064212.1"
/db_xref="GI:28974211"
/translation="MONIINONKHWSEPYKNLYDROVPEKLIINSLKIRHIMYLQGR
RSKSTVFKLIINHLSSKVSKEILYLDNDPFSSTSKDARVINDVLSORLTKR
KVYFLDEISAVNMERYKVSAYDSELPDKIFITSGSALINTQYATLISGRYLTQI
YPLSFKEITLKNININDPELIKERHLYLNVOMAKSGSTVEYTESERKDIISY
YDALILDCVANNQIRIDJAGRELGFSITNFGALISYNSLSNAKLNPKSVQETIC
LQCYLYCDELKMSFYSLSKEQNNKQKYLTDNGLSLAFSDKSGALENLVCEIK
KLGELEFFYKTNSECDPIARKDKSIALOVCELEKNTKREIAGLNKLPFGVNEY
LITFDDRMETELSQFNRLDGLINDNYKHDFPGSDENSETRKSNKLUKIMFYEF
FSERNETYNAL"
complement(110992..14765)
/gene="sap6"
complement(110992..14765)
/gene="sap6"
/note="Cf0006; similar to sap"
/codon_start=1
/transl_table=11
/product="surface array protein"
/protein_id="AA064213.1"
/db_xref="GI:28974212"

[illegible]

Query Match

100.0%; Score 324; DB 6; Length 1271;

Best Local Similarity

100.0%; Pred. No. 5,8e-89;

Matches 324; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCTGTAATTCCTGTTGTCAGATGAGATGACATTAATTTT 60
 DB 853 ATGACAACTGATTTCTCTGTAATTCCTGTTGTCAGATGAGATGACATTAATTTT 912
 QY 61 GCGCTCCGGGCGGCTTCCTGTTAATCCTTAAGCCCTACGTGATGACAAATTTGCGC 120
 DB 913 GCGCTCCGGGCGGCTTCCTGTTAATCCTTAAGCCCTACGTGATGACAAATTTGCGC 972
 QY 121 AAAATGGCGATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 180
 DB 973 AAAATGGCGATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 1032
 QY 181 AGCAATGGCATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 240
 DB 1033 AGCAATGGCATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 1092
 QY 241 GTGGCGCATCTTCTGGCGGTGACGACACCTTGTGAGCGTTGGCGCTGGACATCGTT 300
 DB 1093 GTGGCGCATCTTCTGGCGGTGACGACACCTTGTGAGCGTTGGCGCTGGACATCGTT 1152
 QY 301 TTTGTGACATGATGATGATCTTTTC 324
 DB 1153 TTTGTGACATGATGATGATCTTTTC 1176

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

ORIGIN

BD014994 1271 bp DNA linear PAT 27-AUG-2002
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 BD014994
 BD014994.1 GI:22555801
 JP 2001169788-A/4.
 unclassified.
 ORGANISM
 unclassified.
 1 (bases 1 to 1271)
 Kennerknecht, N., Sahm, H., Eggeling, L. and Pfeiffer, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 2001169788-A 4 26-JUN-2001;
 DEUSSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC13032
 PN JP 2001169788-A/4
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KENNERKNECHT, HERMANN SAHM, LOTHAR EGGELING, WALTER PI
 PFEIFFER
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08//
 PC C12N15/09, C12R1/15, (C12N1/21, C12R1/15), (C12P13/06, C12R1/15),
 PC C12N15/00, (C12N15/00, C12R1/15)
 CC brnF
 CC brnE
 CC brnE
 FH Key
 FT gene
 Location/Qualifiers
 (101)..(853)
 (853)..(1176).
 Location/Qualifiers
 1..1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 Query Match 100.0%; Score 324; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 5,8e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCTGTAATTCCTGTTGTCAGATGAGATGACATTAATTTT 60
 DB 853 ATGACAACTGATTTCTCTGTAATTCCTGTTGTCAGATGAGATGACATTAATTTT 912
 QY 61 GCGCTCCGGGCGGCTTCCTGTTAATCCTTAAGCCCTACGTGATGACAAATTTGCGC 120
 DB 913 GCGCTCCGGGCGGCTTCCTGTTAATCCTTAAGCCCTACGTGATGACAAATTTGCGC 972
 QY 121 AAAATGGCGATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 180
 DB 973 AAAATGGCGATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 1032
 QY 181 AGCAATGGCATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 240
 DB 1033 AGCAATGGCATGATGATGACGAGAGATCTTGCCATTTTGAACGATCAACGTTTGC 1092
 QY 241 GTGGCGCATCTTCTGGCGGTGACGACACCTTGTGAGCGTTGGCGCTGGACATCGTT 300
 DB 1093 GTGGCGCATCTTCTGGCGGTGACGACACCTTGTGAGCGTTGGCGCTGGACATCGTT 1152
 QY 301 TTTGTGACATGATGATGATCTTTTC 324
 DB 1153 TTTGTGACATGATGATGATCTTTTC 1176

RESULT 12

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

gene
 CDS
 Complement (708..1463)
 /gene="brnF"
 Complement (708..1463)


```

RESULT 6
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kemerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched amino
JOURNAL acids, process for the isolation thereof and use thereof
FEATURES
source Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATTTCTCCCTGATTCCTGTTGTCGAGTATGTCAGTCAATTAATT 60
DB 853 ATGACAACGATTTCTCCCTGATTCCTGTTGTCGAGTATGTCAGTCAATTAATT 912
QY 61 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCCCTACGTAATCAATTTGGGCG 120
DB 913 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCCCTACGTAATCAATTTGGGCG 972
QY 121 AAAATGGCGATGTGATGTCAGAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 180
DB 973 AAAATGGCGATGTGATGTCAGAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 1032
QY 181 AGCAATGCGATAGATCTGAAGACTCTAATCCTTGTGATGCGCTTGCGATTACAGTG 240
DB 1033 AGCAATGCGATAGATCTGAAGACTCTAATCCTTGTGATGCGCTTGCGATTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGCGCGGTGTCAGACCACTTGTAGAGTTGGCGTGGACCAATCGTT 300
DB 1093 GTGGCGCATCTTCTTGCGCGGTGTCAGACCACTTGTAGAGTTGGCGTGGACCAATCGTT 1152
QY 301 TTGTTGACTGGTGAATCTTTTC 324
DB 1153 TTGTTGACTGGTGAATCTTTTC 1176

RESULT 7
LOCUS AR391956 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kemerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino
JOURNAL acids, process for the isolation thereof and use thereof
FEATURES
source Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;

```

```

Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATTTCTCCCTGATTCCTGTTGTCGAGTATGTCAGTCAATTAATT 60
DB 853 ATGACAACGATTTCTCCCTGATTCCTGTTGTCGAGTATGTCAGTCAATTAATT 912
QY 61 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCCCTACGTAATCAATTTGGGCG 120
DB 913 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCCCTACGTAATCAATTTGGGCG 972
QY 121 AAAATGGCGATGTGATGTCAGAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 180
DB 973 AAAATGGCGATGTGATGTCAGAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 1032
QY 181 AGCAATGCGATAGATCTGAAGACTCTAATCCTTGTGATGCGCTTGCGATTACAGTG 240
DB 1033 AGCAATGCGATAGATCTGAAGACTCTAATCCTTGTGATGCGCTTGCGATTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGCGCGGTGTCAGACCACTTGTAGAGTTGGCGTGGACCAATCGTT 300
DB 1093 GTGGCGCATCTTCTTGCGCGGTGTCAGACCACTTGTAGAGTTGGCGTGGACCAATCGTT 1152
QY 301 TTGTTGACTGGTGAATCTTTTC 324
DB 1153 TTGTTGACTGGTGAATCTTTTC 1176

RESULT 8
LOCUS AX137709 1271 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1096010.
ACCESSION AX137709
VERSION AX137709.1 GI:14273886
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1 Kemerhnecht,N., Eggeling,L., Sahm,H. and Pfeifferle,W.
AUTHORS Nucleotide sequences coding for branched-chain amino acids export
TITLE Proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 1 02-MAY-2001;
DEGUSA AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source Location/Qualifiers
1..1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/feature="ATCC14752"
101..856
/feature="unnamed protein product; brnF"
/codon_start=1
/transl_table=11
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/db_xref="REMBL:CAC39923"
/translation="MKTQEIHSLSLEVSPSKALPEPDKGRYEIAQKXTSLAQL
GMPIGIAFGILVIQYGEWMAAPLFSGLFAGSTEMLVIALVGAAPLGIATTL
SYWPRVYASFPPIHVKNPPIARVSVFALIDELAVTAAPAMSLIMQIAFH
SWYFGITGTAIHELPIFKIGKEPLKSLFYVLTLDSCRTKQIPSLLAGSFTI
ALVPIPGQALPALLIFGLITIRYFLGRKAK"
853..1179
/feature="unnamed protein product; brnE"
/codon_start=1
/transl_table=11
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMBL:CAC39924"
/translation="MTTDFSCILIVANCAVITPALBAVPLIILKPLRESQFVKAM
MMPGIIAIIITASTFRSNAIDKLITFGLLVAITVAHLIGSRITLISVAGTIVFV
GLVNLFP"

```

[illegible]

RESULT 5	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BD162491	BD162491	Novel polynucleotide.	324 bp	DNA	linear	PAT 17-JAN-2003						
BD162491	BD162491	Novel polynucleotide.	324 bp	DNA	linear	PAT 17-JAN-2003						
BD162491	BD162491.1	GI:27868249										
BD162491	JP 2002191370-A/290.											
BD162491	unidentified											
BD162491	unidentified											
BD162491	unclassified											
BD162491	1 (bases 1 to 324)											
BD162491	Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.											
BD162491	Novel polynucleotide											
BD162491	Patent: JP 2002191370-A 290 09-JUL-2002;											
BD162491	KYOMA HAKKO KOGYO CO LTD											
BD162491	OS Corynebacterium glutamicum											
BD162491	PN JP 2002191370-A/290											
BD162491	PD 09-JUL-2002											
BD162491	PF 15-DEC-2000 JP 2000405096											
BD162491	PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, PI KEIKO OCHIAI,											
BD162491	PI HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENO, MASATO IKEDA, AKIO											
BD162491	PI OZAKI											
BD162491	PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC											
BD162491	C12M1/15,											
BD162491	PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/											
BD162491	PC 04, C12P13/08,											
BD162491	PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC											
BD162491	G01N33/56,											
BD162491	PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1.15),											
BD162491	PC (C12N1/21, C12R1.13), (C12N1/21, C12R1.01), (C12P13/08, C12R1.15),											
BD162491	PC C12N15/00, C12N15/00											
BD162491	PC C12N5/00, C12N15/00											
BD162491	CC Novel polynucleotide											
BD162491	PH key											
BD162491	FT source											
BD162491	FT 1. .324											
BD162491	location/Qualifiers											
BD162491	1. .324											
BD162491	/organism="Corynebacterium glutamicum".											
BD162491	/organism="unidentified"											
BD162491	/mol_type="genomic DNA"											
BD162491	/db_xref="taxon:32644"											
BD162491	ORIGIN											

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 1373.24 seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324

Sequence: 1 atgcgaactgattctctctctg.....cttgactgctgactcttttc 324

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_da.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_da.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_ov.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_hg_hum.*

31: em_hg_inv.*

32: em_hg_other.*

33: em_hg_mus.*

34: em_hg_pin.*

35: em_hg_rtd.*

36: em_hg_mam.*

37: em_hg_vit.*

38: em_sy.*

39: em_hgo_hum.*

40: em_hgo_mus.*

41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	6 AR391955	AR391955 Sequence
2	324	100.0	324	6 AX120374	AX120374 Sequence
3	324	100.0	324	6 AX137712	AX137712 Sequence
4	324	100.0	324	6 BD014993	BD014993 Nucleotide
5	324	100.0	324	6 BD162491	BD162491 Novel pol
6	324	100.0	1271	6 AR391953	AR391953 Sequence
7	324	100.0	1271	6 AR391956	AR391956 Sequence
8	324	100.0	1271	6 AX137709	AX137709 Sequence
9	324	100.0	1271	6 AX137714	AX137714 Sequence
10	324	100.0	1271	6 BD014991	BD014991 Nucleotide
11	324	100.0	1271	6 BD014994	BD014994 Nucleotide
12	324	100.0	2105	1 AF454053	AF454053 Coryneb
13	324	100.0	340000	1 AP005274	AP005274 Coryneb
14	324	100.0	349980	6 AX120085	AX120085 Sequence
15	38.2	11.8	65865	1 AY211269	AY211269 Campylob
16	36.6	11.3	4534	1 CFU25133	CFU25133 Campylob
17	36	11.1	141605	2 AC013732	AC013732 Homo sapi
18	36	11.1	148332	9 AC079586	AC079586 Homo sapi
19	35.8	11.0	2040	8 NT0289862	NT0289862 Nicotiana
20	35.6	11.0	137678	8 AP005183	AP005183 Oryza sat
21	35.6	11.0	173270	9 AC024886	AC024886 Homo sapi
22	35.6	11.0	328187	2 AC117393	AC117393 Homo sapi
23	35.4	10.9	113025	2 AC109596	AC109596 Oryza sat
24	35.2	10.9	34494	3 AC006633	AC006633 Caenorhab
25	34.8	10.7	2000	6 AX655353	AX655353 Sequence
26	34.8	10.7	86524	4 AC104481	AC104481 Sus scro
27	34.8	10.7	270748	2 AC121736	AC121736 Rattus no
28	34.6	10.7	110000	2 AC035209_1	Continuation (2 of
29	34.6	10.7	170475	2 AC136263	AC136263 Rattus no
30	34.6	10.7	179141	2 AC108890	AC108890 Bos tauri
31	34.6	10.7	224691	2 AC137364	AC137364 Rattus no
32	34.6	10.7	270850	2 AC095807	AC095807 Rattus no
33	34.4	10.6	189805	9 AC004169	AC004169 Homo sapi
34	34.4	10.6	190000	2 AC004479	AC004479 Homo sapi
35	34.2	10.6	1686	1 NOCCDH	D90244 Nocardia sp
36	34.2	10.6	12883	1 AE009221	AE009221 Agrobacte
37	34.2	10.6	12909	1 AE008188	AE008188 Agrobacte
38	34.2	10.6	13243	1 AE004602	AE004602 Pseudomon
39	34	10.5	150662	2 AC138550	AC138550 Dario rer
40	34	10.5	167753	9 AC108752	AC108752 Homo sapi
41	34	10.5	168332	2 AC053483	AC053483 Homo sapi
42	34	10.5	169030	2 AC125625	AC125625 Homo sapi
43	34	10.5	251370	5 AL929533	AL929533 Zebrafish
44	33.8	10.4	117645	8 AC144644	AC144644 Medicago
45	33.8	10.4	210791	9 CNS01DV7	AL133467 Human chr

ALIGNMENTS

RESULT 1

AR391955

LOCUS AR391955 324 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 4 from patent US 6613545.

ACCESSION AR391955

VERSION AR391955.1 GI:40115726

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 324)

AUTHORS Kennerknecht, N., Sahm, H., Eggeling, J., and Pfeifferle, W.

TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof

JOURNAL Patent: US 6613545-A 4 02-SEP-2003;

Matches 108; Conservative 0; Mismatches 92; Indels 6; Gaps 1.

Oy 223 TTGCGGGCTCCACCGAATGCTGTCATGCCCTCGTTGTGGGCGCAAGGCCCTGGGC 282

Matches 108; Conservative 0; Mismatches 92; Indels 6; Gaps 1;

Db	278	TTTGCGCGATCCACCGAGCGGCGCTCATCTCTCTTGGGCGCGCTCGGCGCGTGGCGAAC	219
Qy	283	GGCACCAGCGCTCACCAATGTCGTCGAACTTGCAGCAGTAATTCATAGCGTTTTCATTC	342
Db	218	GGCGCGCGATGCCCTTCATGACCGGTCTTTGCGCACTCATCGACCGCTTTGCGGCGC	159
Qy	343	CGCGT-----CATGTGTCAAAACCCCATTTGCCGTTTCTATTGCGTTTCGCGCTT	396
Db	158	GGGCTCAGCATCCACGAGTGTGTAAACGGGTGTCCATGGTCTCTCGATTTCGTAC	99
Qy	397	ATCGACGAAGCTTACGCGAGTCACTGC	422
Db	98	CTGCGCATCGCCTCGCGCGTGCCTTC	73

LOCUS CDS0165R 1201 bp DNA linear GSS 26-JUL-1993
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BAC13514.4 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL106545
AL106545.1 GI:5622456
VERSION 1
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyridiidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
cosmoseq. cosmid library
cosmoseq. cosmid library

Submitted (23-09-2019) to Genoscope - Centre national de séquençage
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
- Determination of this BAC-end sequence was carried out as part of
collaboration of the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBeloBAC11.

```
source
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15K14"
/clone_lib="DrosBAC"
```


Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 1 (bases 1 to 697)
 Myler,P.J., Vogt,C., Munden,H., Robertson,L., Sisk,E., Fazelina,G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E., Stuart,K. and Ragland,M.
 Leishmania major Friedlin BAC End Sequences
 Unpublished (2002)
 Other GSSs: LB01712a.d_17.1
 Contact: Myler PJ
 Seattle Biomedical Research Institute
 4 Nickerson Street, Seattle, WA 98109-1651, USA
 Tel: 206 284-8846
 Fax: 206 284-0313
 Email: mylerpj@brl.org
 Seq primer: Sp6
 Class: BAC ends.
 Location/Qualifiers
 1..697
 /organism="Leishmania major"
 /mol_type="genomic DNA"
 /strain="Friedlin"
 /db_xref="taxon:5664"
 /clone="LB01712a"
 /lab_host="E. coli GeneHogs + TrfA"
 /clone_lib="Leishmania major Friedlin BAC Library"
 /notes="Vector: pCG270; Site 1: HindIII; Genomic DNA from Leishmania major Friedlin in agarose blocks was partially digested with HindIII, size selected, and ligated with HindIII-digested pCG270 vector DNA. 10368 clones were picked and arrayed in 384- and 96-well plates. Library construction and arraying was carried out by ResGen Corporation and clones and filters are available from them"
 ORIGIN
 Query Match 5.0%; Score 37.4; DB 28; Length 697;
 Bacc Local Similarity 52.2%; Pred. No. 46;
 Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 Oy 181 GGCTACCAATGTTGGGAGCCCACTGTTTCCGGCCTGATTTTCGGGGCTCCACCGAA 240
 Db 579 GGCAGCCCATGTCGTGTCGCGACACTCCCGACCCACATCGCTTCGCGCGGCG 520
 Oy 241 ATGCTGTCATCGCCCTCTGTTGTGGGGCGACGCCCTCGGGCGCCATCGGCTACACAA 300
 Db 519 CACGAGTCGGCGCGCGCTGCGAGCCTCTGTCGCGCTTGGCGGCAACCCGTTACCGAAG 460
 Oy 301 TTGCTGGGAACCTTCGCCACGATTCATGCGTTTCA 339
 Db 459 CCACTGTCACCTGTCATGTCCTCGTCCGACATCTCA 421
 RESULT 10
 LOCUS BX382084/c 1108 bp mRNA linear EST 08-MAY-2003
 DEFINITION BX382084 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1074Yc24 5-PRIME, mRNA sequence.
 ACCESSION BX382084
 VERSION BX382084.1 GI:30455015
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1108)
 Li,W.B., Gruber,C., Jeesee,J. and Playas,P.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segrret@genoscope.cns.fr, Web : www.genoscope.cns.fr

Db		1100	SCCMMBGSKGGCKKKKMCBCCMBCBKVCYVKKKKCKKKKKCKKKCKNNKK	1159
Cy		599 GC	600	
		:	!	
Db		1160 BC	1161	
RESULT 6				
CNS0091P				
LOCUS				
DEFINITION				
	CNS0091P	925 bp	DNA	linear GSS 03-JUN-1999
	Drosophila melanogaster genome survey sequence TERT3 end of BAC #			
	BACR1D16 of RPCT-98 library from Drosophila melanogaster (fruit			
	fly), genomic survey sequence.			
	AL053013			
	AL053013.1	GI:4934461		
	GSS.			
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
	Drosophila melanogaster (fruit fly)			
	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
	Ephydroidea; Drosophilidae; Drosophila.			
	1 (bases 1 to 925)			
	Genoscope.			
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
	Direct Submision			
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :			
	Pathol Biol 1996; 42(1): 1-7			
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : cna.fr			
	Pathol Biol 1996; 42(1): 1-7			

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kaetucyo Osoegawa and Aaron Mammaster in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACpac Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
  source
    location/Qualifiers
    1..925
    /organism="Prosochila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BAC19D16"
    /clone_lib="RFCl-96"
    /note="end : TET3"

```

ORIGIN

Query Match 5.2%; Score 39; DB 29; Length 925

Best Local Similarity 14.2%; Pred. No. 19;
Matches 37; Conservative 117; Mismatches 106; Indels 0; Gaps 0

195 GCGACGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCAATGCTGTCATCGC 254

Db 638 SSSSSSTSSSTSSSTKSSSSGSSSSSSSSTTSKSTASAGSGSWAGGSGSTGSTSS 697

255 CCTGTGTGTGGGCGGAGCGCCCTGGGCGCCATGCGGCTCACCACATTGGTGTGAATT 314

Db 698 SSSSSSTSTSSSVSSGKSLSTBSSGGBSSSSSSSTSSBBSCTSTSSSSSSSYSSST 757

[illegible]

375 TTCTATTGGTTTTCGCGCTTATCGACGAAGCTACGCACTCCTCGGCGCAGGCCCGC 434

818 YMBCYSTSCGSSSSSGKGYTKCGCGCGSSSTNMBGTSSACSSSSSSCSSSSVSS 877

```

QY      435 AGGCTGGTCGGCGTGGCGAC 454
          ::  : :::: |::|::
Db      878 SSKSSASSSVSSSGSSGVS 897

```

RESULT 7	EX361080	LOCUS	DEFINITION
	EX361080	1201 bp	mRNA linear
	EX361080	Homo sapiens PLACENTA COT 25-NORMALIZED	Homo sapiens cDNA

ACCESSION	GI001079110.5-FRAME, mRNA sequence.
VERSION	BX361080
KEYWORDS	BX361080.1 GI:30374504
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope
	Genoscope, Centre National de Séquençage

Genoscope - Centre National de Séquençage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
<http://fulllength.invitrogen.com> Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS001079DGG08NP1.

source	location/qualifiers
1. .1201	

```

//organism="Homo sapiens"
//mol_type="mRNA"
//db_xref="taxon:9606"
//clone="CSOD10791N16"
//tissue_type="PLACENTA COT 25-NORMALIZED"
//clone_tlib="Homo sapiens PLACENTA COT 25-NORMALIZED"
//note="1st strand cDNA was primed with a NotI/ligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	5.14; Score 38.6; DB 13; Length 1201;
-------------	---------------------------------------

Best Local Similarity 10.5%; Pred. No. 28;
Matches 34; Conservative 147; Mismatches 142; Indels 0; Gaps 0

19 ATTCAATTCAGCCTCGAGGTGTGGCATCCAGGACGCCCTGAAACAGATGATAAGGT 78

825 MDDMMNMGINDKKKKNIDKKGMKKAKKKKKKKD KKKMDKD KMMKKKKKAMKKAWAKKK 884

79 TATCGGCGCTACGAATTCGGCCAGGTCTAAAACTCCCTTGCATGAGGTTGGGCATG 138

+: :: : |::: : : ::::: +::: :

[illegible]

Dy	I39	TACCGATTCGTTGGTGGCTCCTCGATTATCCAAAGGCACGGAATGCAAGGACA	178
	:	: : :	:
Dh	945	KKKKKKKKTGTTTTTTTTTTTMMNNNNNNTTTTTTTTTTTTTHNKKKKKKKKKK	1000

199 GCGCCACTGTTTTCGGCCCTCAATTTTCGGCGGCTCCACCGAATGCTGCTCATCGCCCTC 258

Db 1.005 KADNDMMMMMKKKKMHKKKKKKKKKKMMKKKKKKMMKKKKKKHNNKNNNNNNKKKKKKMMNNNNKK 106

259 GTTGTGGAGCGCAGCGCCCGTGGAGCGCATCGGCTACACCAATGTCTGGTGAATTCCGC 318

```
Db      1065 KKKKKKCCCCBEBCCCCCKYKCKBKXCCYKCKKKKKCKKKYCCCBCCCCKKKCBYCC 112
```

319 CACGATTTCTANGCGTTTTCATT 341

Db	1125	CCCCCCCCCCCCCCCC	1147

8. RESULT

OY		491	TATTCGGGCTCACCGAGTGGCATTCGACAGTTGATTCTTTGAATTAAAGGCC	550
Dd		535	GTTTCGGCGCCGCACCACCGAATCAACTTGAGCTTCGCGAGCGCCTTAAGAATTCAAGTC	594
OY		551	TCGAGTTCCGCTTGTCTCTCTCTTTGTTCAGCG	583
Dd		595	ACGAGAGGCGCTTGTCTCTCTCTCTCTCTCCGCGC	627
RESULT 4	CNS009MA/c			
LOCUS				
DEFINITION	CNS009MA	787 bp	DNA	linear GSS 03-JUN-1999
VERSION	Drosophila melanogaster genome survey sequence T7 end of BAC #			
KEYWORDS	BACR20K04 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
SOURCE	AL054013	GI:4935590		
ORGANISM	GSS. Drosophila melanogaster (fruit fly)			
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 787)			
AUTHORS	Genoscope.			
TITLE	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr			
JOURNAL	- Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mamoser in pletier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.			
FEATURES	Location/Qualifiers			
source	1..787 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACR20K04" /clone_1lb="RPCL-98" /note="end : T7"			
ORIGIN				
Query Match	5.2%; Score 39.4; DB 29; Length 787;			
Best Local Similarity	45.5%; Pred. No. 14;			
Matches	76; Conservative 15; Mismatches 76; Indels 0; Gaps 0			
OY		552	CGAGTTGCGCCCTTGTCTCTCTTTGTTCACGCTGATTTGATTTCTTGCAGAACAAAA	611
Dd		334	CTATCTTCGTCCTTCTTCTTTTMTTATAAKMKMCTGGAAATCKRTVAGARPAAC	275
OY		612	GCAGATCCCTTCTGTGTCTCGACAGTTGAGCTTCACCATTTGCTTGTGGTAATCC	671
Dd		274	AATTTGTTTTTKTTTCYACWGRCTTAGAGTAATTCTTAATTTTWTGTTTAATTT	215
OY		672	AGGTACAGCCCTATTGTGGCGCTGTGATCTTTGGGTTGTTGA	718
Dd		214	TAAATTCGCATATGATTATATCGAATTTTATTATGTCGCGTA	168
RESULT 5				
EX381961				

LOCUS	BX381961	1201 bp	mRNA	linear	EST 08-MAY-2003
DEFINITION	BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSOD1072YF05 3-PRIME, mRNA sequence.				
ACCESSION	BX381961				
VERSION	BX381961.1	GI:30453007			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001)				
AUTHORS	Contact: Genoscope				
TITLE	Genoscope - Centre National de Sequencage				
JOURNAL	BP 191 91006 EVRY cedex - France				
COMMENT	Email: seqest@genoscope.cns.fr Library was constructed by life technologies, a division of invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitroden Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOD1072CC03NP1.				
FEATURES	location/Qualifiers				
Source	1..1201				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSOD1072YF05"				
	/issue_type="PLACENTA COT 25-NORMALIZED"				
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"				
	/note="1st strand cDNA was primed with a NotI-clig0 (dr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pcwvSPORT 6 vector. Library was normalized."				
ORIGIN					
Query Match	5.2%; Score 39.2; DB 13; Length 1201;				
Beef Local Similarity	6.3%; Pred No. 19;				
Matches	34; Conservative 210; Mismatches 299; Indels 0; Gaps 0;				
Db	59 TGAACGATGATAAGATTATCGGGCGCTACGAATCGCGAAGTCTAAAACTCCC 118				
	:::::~::~:				
Db	620 KKKKKNMNNNKK 679				
Qy	119 TTGCAGCAGTTTGCGCATGTACCAGATTGATGATTGTCCTTGGTATTCAAT 178				
	:::~::~:				
Db	660 KKKKKNN 739				
Qy	179 ACGGCTACGAATGTGGCAGCCCCACTGTTTTCCGGCTAATTTCCGGGCTCACCG 238				
	:::~::~:				
Db	740 MNNNCNCKKKNNNTNMNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCK 799				
Qy	239 AAATCGTGTATGCGCCTCGTTGGGGCGCAGCGCCCTGGGGCGCATCGCGCTCAC 298				
	:::~::~:				
Db	800 MCCMKKKKKKKKKKKKKCMCKMMNNKKKCMCKBMMCKMKCKMBKNNMMMMMKM 859				
Qy	299 CATTCGTGGTAACCTTCGCCGACGATTTCTATGCGTTTCATTCGGCTGATGGTCA 358				
	:::~::~:				
Db	860 MNN 919				
Qy	359 AAAACCCCATTCGCCGTTCTATTGGGTTTTGGCGCTTAGAGAAGCTACGAGTCA 418				
	:::~::~:				
Db	920 MMCKKBMMCKKNNN 979				
Qy	419 CTGCGCGCAGGCGCCGAGCTGTCGGCGGTGGCGACTTATCTCAATGACAATAGCTTT 478				
	:::~::~:				
Db	980 VKKVMCMDBKNNCKKKBMKVCCCKKCMNNNNNNNNNNNNNNNNNNNNNNNNNN 103				
Qy	479 ACTCTACTGGGTATTCCGCGGTCTCACCGAAGTGGCGATGCGAAGTTATCTTTTG 538				
	:::~::~:				
Db	1040 MNN 1099				
Qy	539 AAATTAAAGGCTCGAGTTGCGCCTTGTCTCTTTTGACGCTGACTTTGATTCCT 598				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 2438.38 seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753
Sequence: 1 gtgcacaaacgaagagat.....ctctgggaaagctgctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

BST:
1: em_eebba:*
2: em_eebba:*
3: em_eebba:*
4: em_eebba:*
5: em_eebba:*
6: em_eebba:*
7: em_eebba:*
8: em_eebba:*
9: gb_eebba:*
10: gb_eebba:*
11: gb_eebba:*
12: gb_eebba:*
13: gb_eebba:*
14: gb_eebba:*
15: em_eebba:*
16: em_eebba:*
17: em_eebba:*
18: em_eebba:*
19: em_eebba:*
20: em_eebba:*
21: em_eebba:*
22: em_eebba:*
23: em_eebba:*
24: em_eebba:*
25: em_eebba:*
26: em_eebba:*
27: em_eebba:*
28: gb_eebba:*
29: gb_eebba:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54.6	7.3	473	28	AQ936633 HSJ41-905
2	45	6.0	1201	13	EX381961 BX381961
3	39.4	5.2	743	29	CC684922 OGUAX55TH
4	39.4	5.2	787	29	CNS009WA

5	39.2	5.2	1201	13	EX381961	EX381961 BX381961
6	39	5.2	925	29	CNS0091P	AL050303 Drosophila
7	38.6	5.1	1201	13	EX381961	EX381961 BX381961
8	38	5.0	835	14	CB657749	CB657749 OJNEC13F
9	37.4	5.0	697	28	BH887860	BH887860 LB01712A
10	37.4	5.0	1108	13	EX382084	EX382084 BX382084
11	37.2	4.9	730	12	BG845919	BG845919 1024011FO
12	37.2	4.9	1454	12	BF101158	BF101158 601754733
13	36.8	4.9	464	28	BH805539	BH805539 1008061A1
14	36.8	4.9	1201	13	CNS016BR	AL050303 Drosophila
15	36.6	4.9	539	10	BH013856	BH013856 125354 MA
16	36.6	4.9	942	29	CNS000CT	AL0509804 Drosophila
17	36.4	4.8	298	12	BH525053	BH525053 B1525053
18	36.2	4.8	673	12	BH287010	BH287010 B1787010
19	36.2	4.8	842	13	EX381708	EX381708 BX381708
20	36.2	4.8	902	29	CG358991	CG358991 OGA3A035TH
21	36.2	4.8	1101	13	BG678281	BG678281 AGENCOURT
22	36.2	4.8	559	13	EX384405	EX384405 BX384405
23	36	4.8	666	28	CG170529	CG170529 1197d07.b
24	36	4.8	784	29	CG250542	CG250542 OGUH070TV
25	36	4.8	836	29	CG449984	CG449984 OGUH070TV
26	36	4.8	852	28	BZ644270	BZ644270 OGUH070TV
27	36	4.8	863	29	CG449978	CG449978 OGUH070TV
28	35.8	4.8	410	9	AI138160	AI138160 TUES723 T
29	35.8	4.8	641	28	BZ375719	BZ375719 160910.9
30	35.8	4.8	641	29	CG668174	CG668174 OGUH070TV
31	35.8	4.8	696	29	CG654573	CG654573 OGUH070TV
32	35.8	4.8	980	29	CG326716	CG326716 OGUH070TV
33	35.6	4.7	327	10	BF543765	BF543765 UI-R-C3-C
34	35.6	4.7	515	13	BK424977	BK424977 BX424977
35	35.6	4.7	592	9	AI293892	AI293892 LP07146.5
36	35.6	4.7	657	13	CA109167	CA109167 SCGRH106
37	35.6	4.7	683	14	CP850119	CP850119 PMA013XA
38	35.6	4.7	684	28	CG337836	CG337836 OGUH070TV
39	35.6	4.7	798	12	BG855128	BG855128 1024011FO
40	35.6	4.7	900	29	CG371461	CG371461 OGUH070TV
41	35.6	4.7	977	29	CNS000CT	AL076880 Drosophila
42	35.6	4.7	985	13	BK359916	BK359916 BX359916
43	35.6	4.7	1201	13	EX381971	EX381971 BX381971
44	35.4	4.7	398	14	CP65113	CP65113 PZ8014XB
45	35.4	4.7	454	9	AI575130	AI575130 UI-R-GO-U

ALIGNMENTS

RESULT 1
LOCUS HSJ41-905 Human Ncti clones Homo sapiens genomic, genomic survey
DEFINITION
ACCESSION AQ936633
VERSION AQ936633.1 GI:7213011
KEYWORDS
SOURCE GSS
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Zabarovsky E.R., Gazatullin R., Podowski R.M., Zabarovsky V.V.,
Mammalia Euthetia, Primates, Catarrhini, Homiidae, Homo.
1 (bases 1 to 473)
Li, J., Protodopov, A., Kashuba, V., Ernerberg, I., Winberg, G. and
Mahlstedt, C.
Ncti clones in the analysis of the human genome
Nucleic Acids Res. 28 (7), 1635-1639 (2000)
TITLE
JOURNAL MEDLINE
PUBMED
COMMENT
Contact: Podowski R.M.
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337983

Tue Apr 20 06:47:25 2004

us-10-608-504-6_copy_101_853.rnpn

Page 6

Search completed: April 17, 2004, 20:13:05
Job time : 88.0151 secs

RESULT 12

US-60-548-091-19934
; Sequence 19934, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19934
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19934

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 301
DB 46 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 105
QY 302 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 361
DB 106 GCTTACAGGACACCGGCAATTTCACTCATGCGTCCGCTGCGGATTCTACC 165
QY 362 ACCCATTCGCC 372
DB 166 TCGCCCTGACC 176

RESULT 13

US-60-548-091-19938
; Sequence 19938, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19938
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19938

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 301
DB 49 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 108
QY 302 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 361
DB 109 GCTTACAGGACACCGGCAATTTCACTCATGCGTCCGCTGCGGATTCTACC 168
QY 362 ACCCATTCGCC 372
DB 169 TCGCCCTGACC 179

RESULT 14
US-10-767-701-9054

; Sequence 9054, Application US/10767701

; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9054
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1
US-10-767-701-9054

Query Match 4.4%; Score 32.8; DB 6; Length 1028;
Best Local Similarity 54.0%; Pred. No. 1.7;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 609 AAGCAGATCCCTTCTCTGCTGCTCGCAGGTTGACCTTCAACATGCTTGTGTAN 668
DB 746 ATAGCCCTTCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
QY 669 TCCAGGTACAGCCCTATTTCGCGCGCTGATCTTCTGAGTCTGTCAGCATCCGTA 728
DB 806 GCCCAATATACCCCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
QY 729 CTTC 732
DB 866 GTCC 869

RESULT 15

US-10-767-701-23189
; Sequence 23189, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 23189
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14593439
US-10-767-701-23189

Query Match 4.3%; Score 32.6; DB 6; Length 519;
Best Local Similarity 52.6%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 182 GCTACGATGTGTGGAGAGCCCACTGTTTCCGCTGATTTTCCGCGCTCCACCGAAA 241
DB 70 GCTACGATGTGTGGAGAGCCCACTGTTTCCGCTGATTTTCCGCGCTCCACCG 129
QY 242 TCGTGTATCGCCCTCGTTGTGGCGAGCGCCCTGCGCCATCGCGTCAACCAT 301
DB 130 CCTGTGTGCGCGCTGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 189
QY 302 TCGTGTATCGCC 316
DB 190 CCGTGTCAAGTACC 204

Query Match 4.5%; Score 34; DB 6; Length 561;
Best Local Similarity 54.9%; Pred. No. 0.53;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 202 CCACGCTTTTCCGGCTGATTTTCCGGGCTCCACCAATGCTGATGCGCTGCT 261
DB 333 CCACGAGCGGCGACCTTGAGCTTTGAGCGGAGCGGATCTGGGCTGCTGCTC 274
QY 262 GTGGCGGAGCGGCTGGGCGGCGATCGGCTGACACATTTGCTGTAAGTTCGCGAC 321
DB 273 ACCAGCGCGAGCGGCTGGGCTGCTTCTCTCCACACCTTGTGATGAGTACTGCGCC 214
QY 322 GT 323
DB 213 CT 212

RESULT 9
US-10-765-790-62
Sequence 62, Application US/10765790

GENERAL INFORMATION:
APPLICANT: Bayer Healthcare LLC
APPLICANT: Beard, Chris
APPLICANT: Burgess, Chris
APPLICANT: Gannon, Allison
APPLICANT: Harvey, Jeanne
APPLICANT: Lechner, John F.
APPLICANT: Li, Zheng
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
FILE REFERENCE: 1657/2035
CURRENT APPLICATION NUMBER: US/10/765,790
PRIOR FILING DATE: 2004-01-27
PRIOR APPLICATION NUMBER: US 10/737,082
PRIOR FILING DATE: 2003-12-16
NUMBER OF SEQ ID NOS: 300
SOFTWARE: PatentIn version 3.2
SEQ ID NO 62
LENGTH: 23490
TYPE: DNA
ORGANISM: Homo sapiens
US-10-765-790-62

Query Match 4.5%; Score 33.8; DB 6; Length 23490;
Best Local Similarity 49.7%; Pred. No. 3.8;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 201 CCCGCTTTTCCGGCTGATTTTCCGGGCTCCACCAATGCTGATGCGCTGCT 260
DB 21210 CTCGCTCTCTTCCCTGCTGATCTGCTGCTCACTTCTGCTGCTGAGGCTGGGA 21269
QY 261 TGTGGGCGAGCGGCTGGGCGGCGATCGGCTGACACATTTGCTGTAAGTTCGCGCA 320
DB 21270 GGAAGCTGGGCGGCTGGGCTGGGCTGGGCGGCTGATGCTGCTGCTGAGGCTGACT 21329
QY 321 CGATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
DB 21330 CCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21382

RESULT 10
US-60-548-091-5686
Sequence 5686, Application US/60548091
GENERAL INFORMATION:
APPLICANT: Carlgill, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CLO01506
CURRENT APPLICATION NUMBER: US/60/548,091
CURRENT FILING DATE: 2004-02-27
NUMBER OF SEQ ID NOS: 24433
SOFTWARE: fastseq for Windows Version 4.0
SEQ ID NO 5686

LENGTH: 24477
TYPE: DNA
ORGANISM: Homo sapiens
US-60-548-091-5686

Query Match 4.5%; Score 33.8; DB 7; Length 24477;
Best Local Similarity 51.9%; Pred. No. 3.9;
Matches 68; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 242 TGGTGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
DB 17603 TGGTGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 17662
QY 302 TGGTGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
DB 17663 GCTTCAAGAGCAGCGGCGGCGGATTTCACTCCATGAGCGTTCGCGCTGCGGATTTACC 17722
QY 362 ACCCATTTGCC 372
DB 17723 TGCCCTGGCC 17733

RESULT 11
US-09-804-291A-500
Sequence 500, Application US/09804291A

GENERAL INFORMATION:
APPLICANT: ZOZULA, SERGEY
TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 100337,54287US
CURRENT APPLICATION NUMBER: US/09/804,291A
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/188,914
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/192,033
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/198,474
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/199,335
PRIOR FILING DATE: 2000-04-24
PRIOR APPLICATION NUMBER: 60/207,702
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/213,849
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/226,534
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/230,732
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/266,862
PRIOR FILING DATE: 2001-02-07
NUMBER OF SEQ ID NOS: 531
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 500
LENGTH: 930
TYPE: DNA
ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match 4.4%; Score 33.2; DB 5; Length 930;
Best Local Similarity 54.0%; Pred. No. 1.2;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 467 AAATAGCTTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526
DB 509 AATTTCTGATCAGCATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568
QY 527 TGAATCTTTGAAATTAAGGCGCTGAGTTGCGCTTGTCTCTCTTTGTACAGCTGA 586
DB 569 CAGGCTCTTAAGATGATGCTCATCTTGTGCGGCTGTGATCTCATGTTACTTGG 628
QY 587 CTTTGG 592
DB 629 TGCTGG 634

```
/ Sequence 983, Application US/10100683
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et al.
/ TITLE OF INVENTION: Human Secreted Proteins
/ FILE REFERENCE: PS900
/ CURRENT APPLICATION NUMBER: US/10/100,683
/ PRIOR FILING DATE: 2002-03-19
/ PRIOR APPLICATION NUMBER: US 60/040,162
/ PRIOR FILING DATE: 1997-03-07
/ PRIOR APPLICATION NUMBER: US 60/043,576
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,601
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,845
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,580
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,599
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,664
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,314
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,632
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,892
/ PRIOR FILING DATE: 1997-08-22
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 13468
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 983
/ LENGTH: 3173
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-100-683-983

Query Match          4.6%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.79;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGTCATCGCCCTCGTGTGGGCGAGGCGCCCTGGGGCGCATGCGCTCACACCA 300
DB 2360 CAGCTCCCGGCCCGCCAGTGGGCTCAGGCGCTCTGGGCGCCATGGCTCTCACCA 2304

RESULT 6
US-10-100-683-11511
/ Sequence 11511, Application US/10100683
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et al.
/ TITLE OF INVENTION: Human Secreted Proteins
/ FILE REFERENCE: PS900
/ CURRENT APPLICATION NUMBER: US/10/100,683
/ PRIOR FILING DATE: 2002-03-19
/ PRIOR APPLICATION NUMBER: US 60/040,162
/ PRIOR FILING DATE: 1997-03-07
/ PRIOR APPLICATION NUMBER: US 60/043,576
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,601
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,845
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,580
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,599
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,664
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,314
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,632
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,892
/ PRIOR FILING DATE: 1997-08-22
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 13468
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 983
/ LENGTH: 3173
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-100-683-983
```

```
/ PRIOR FILING DATE: 1997-08-22
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 13468
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 11511
/ LENGTH: 7071
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-100-683-11511

Query Match          4.6%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.2;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGTCATCGCCCTCGTGTGGGCGAGGCGCCCTGGGGCGCATGCGCTCACACCA 300
DB 785 CAGCTCCCGGCCCGCCAGTGGGCTCAGGCGCTCTGGGCGCCATGGGTCTCACCA 841

RESULT 7
US-10-767-701-7545
/ Sequence 7545, Application US/10767701
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5335)B
/ CURRENT APPLICATION NUMBER: US/10/767,701
/ PRIOR FILING DATE: 2004-01-29
/ NUMBER OF SEQ ID NOS: 63128
/ SEQ ID NO 7545
/ LENGTH: 666
/ TYPE: DNA
/ ORGANISM: Sorghum bicolor
/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
/ US-10-767-701-7545

Query Match          4.6%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.42;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 215 GCTGATTTTTCGCGGCTCCACCGAATGCTGCTATCGCCCTGTGTGGGCGCACGCG 274
DB 31 GCCGACCAACCAACGAGGAGGACGACCCGCTCATCAACCGTTGTGACGCGCGCGCG 90

QY 275 CCTGGGCGCCATCGCGCTCACCATTTGCTGTGAACCTTCGCCACG 322
DB 91 CGCTGCGCTCCCTGCGTGGCGAGCGCGGTGCTGTGTGGTGGCGAG 138

RESULT 8
US-10-767-701-744/c
/ Sequence 744, Application US/10767701
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5335)B
/ CURRENT APPLICATION NUMBER: US/10/767,701
/ PRIOR FILING DATE: 2004-01-29
/ NUMBER OF SEQ ID NOS: 63128
/ SEQ ID NO 744
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Sorghum bicolor
/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
/ US-10-767-701-744
```


Page 2

```

APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5263
LENGTH: 619
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS11974_1
US-10-767-701-5263

Query Match          4.7%; Score 35.4; DB 6; Length 619;
Best Local Similarity 54.1%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 61;

Dy      226 GCGGGCTCCACCGAATGCTGTCATGCGCCCTCTGTGTGGCGCAGCGCCCTGGGCGCC 285
      |||||
Db      132 GCGGCTTACCTGCTGCTGCTGCTGCGGGCTGGCGGTGACGGGCAAGTGCTGGGCTC 191

Dy      286 ATGCGCTACACATTTCTGTGTAACCTCCGCGACGATTTCTATGCGTTTCATCCCG 345
      |||||
Db      192 ATGTGGCAGCGCGCTGCTGTGATTTTCAGCCCGCTGCTGTGCGCGGCATCAGC 251

Dy      346 CTGCATGTGTCA 358
      |||||
Db      252 GTGGCGCTGCTCA 264

RESULT 4
US-10-767-701-10866
Sequence 10866, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kowalec, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 10866
LENGTH: 1034
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
US-10-767-701-10866

Query Match          4.6%; Score 34.6; DB 6; Length 1034;
Best Local Similarity 49.7%; Pred. No. 0.45; Indels 0; Gaps 0;
Matches 88; Conservative 0; Mismatches 89;

Dy      241 ATGTGTATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGCATGCGCTGACGACA 300
      |||||
Db      151 ATGGCGCGCCCTTACTACGCGGAGGCCCTGTGCGCCGACACGGCGGGCGGCCCT 210

Dy      301 TTGCTGTGAATTCGCGCCACGATTTCTATGCGTTTCTATTCGCGCTGCAATGTGTCAA 360
      |||||
Db      211 GTGCGGTGTGTGTCGCCGCGAGTTCTGTGCGCGCGTACGTGTGCCCTTACGGTACCAAG 270

Dy      361 AACCCATTCGCGTTTCTATTCGTTTGTGGGCTTATGACGAGAGCTTACGACATG 417
      |||||
Db      271 AAGGCATACGCTCTTCGACGGCGCACTTACCGTACCCGACGACGAGGCGGCGCTC 327

```


APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(751)
OTHER INFORMATION: n = A,T,C or G
US-09-439-313-12

Query Match 4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred.No.0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 227 CGGGCTCCACCGAATGCTGTCATGCGCCCTGTTGTGGCGCAGCGCCCTGGGGCCCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTCACCAATGCTGCTGTAAGTTCCGCGCAGCT 323
DB 350 TGGTGTACACCAATGCTGAGCACTTCCCTGAGCT 314

RESULT 15

US-09-352-616A-12/C
Sequence 12, Application US/09352616A
Patent No. 6395278
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Wang, Yugu
APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer Lynn
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.427C8
CURRENT APPLICATION NUMBER: US/09/352,616A
CURRENT FILING DATE: 1999-07-13
NUMBER OF SEQ ID NOS: 472
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(751)
OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-12

Query Match 4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred.No.0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 227 CGGGCTCCACCGAATGCTGTCATGCGCCCTGTTGTGGCGCAGCGCCCTGGGGCCCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTCACCAATGCTGCTGTAAGTTCCGCGCAGCT 323
DB 350 TGGTGTACACCAATGCTGAGCACTTCCCTGAGCT 314

FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-9300
 TELEFAX: (703) 683-4109
 TELEAX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZgpc-Fls
 US-08-232-463-14

Query Match 6.1%, Score 46.2; DB 1; Length 7218;
 Best Local Similarity 5.2%; Pred. No. 0.00027;
 Matches 21; Conservative 214; Mismatches 172; Indels 0; Gaps 0;

215 GCTGATTTTCGCGGCTCCACGAAATGCTGATCGCCCTGTTGTCGCGACGC 274
 1041 GGCTGAGGTGAGGAGCTTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1100
 275 CCTGGGGGCGCATGCGCTCACACATTGCTGTAATTCGCCGCTATTCATGCT 334
 1101 TT 1160
 335 TTATATCCCGCTCATGCTGTAATAAACCAGTCCGCTTCTATGCTTTTCGCG 394
 1161 TT 1220
 395 TTATCGAGAGCTAGCAGCTCAGTGGGCGGCGGCGGAGGCTGCGCTGCGAC 454
 1221 TT 1280
 455 TTATCATGCAATAGCGTTTCACTCTACTGCGGATTCGCGGCTCACCGAGTGG 514
 1281 TT 1340
 515 CGATCGAGAGTATTCCTTTGAATTAAGGCGCTCGAGTTCGCGCTTGTCTCT 574
 1341 TT 1400
 575 TTGTCAGCGCTGCTTGGATTCCTCGCAAGGAGGAGATCCCT 621
 1401 TT 1447

RESULT 6
 US-09-107-532A-1266
 Sequence 1266, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7210
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-09-107-532A-1266

Query Match 5.3%; Score 39.8; DB 4; Length 768;
 Best Local Similarity 4.93%; Pred. No. 0.0092;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

107 TAAACCTCCCTGCTGAGGTTTGCGCATGACCCGATTTGCGGCTGATTTCT 166
 92 TCAAGACAGCTTACCTACCGTTTGGTTATATGATTTGACTTCATTGTATGAC 151
 167 TGGTATTCAATAGCGTACGAATGATGAGGAGCCCACTGTTTCCGCGCTGATTTG 226
 152 TTGGGAAGGTGCGGATTTTCACTAGTGTGTCACGTTGATGCTGCGTGTATG 211
 227 CGGCTCACCGAAGTGTGATGCTGCTGTTGCGGCGAGCGCCCTGCGGCGCA 286
 212 CTGTTCTGCGCAATTTATACAGTACGATGCTTGTAGTACAGAGCCCATTTGCA 271
 287 TCGGCTCACCATTTGCTGTAATTCGG 317
 272 TCGTTTCTGACCTTTCTAGTCAATTCGG 302

RESULT 7
 US-09-252-991A-11376
 Sequence 11376, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 11376
 LENGTH: 1038
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa

QY	GTGCAAAAAACGCAAGATTATTCACAGCCGTGAGAGGTGTGGCCATCAAGAGCCCTG	60
Db	GTGCAAAAAACGCAAGATTATTCACAGCCGTGAGAGGTGTGGCCATCAAGAGCCCTG	160
QY	GAAACAAGATATAAGATTATCGGCGCTACGAAATCGCGCAAGCTCTAAAACTCCCTT	120
Db	GAACCAAGATATAAGATTATCGGCGCTACGAAATCGCGCAAGCTCTAAAACTCCCTT	220
QY	GCACGAGTTTGGGCAATGACCCGAATGGTAATGGGTTTGGTCTTGGTATTCAATAC	180
Db	GTGCAAGTTTGGGCAATGACCCGAATGGTAATGGGTTTGGTCTTGGTATTCAATAC	280
QY	GACTACGAATGATGGGCAAGCCCACTGTTTCCGACCTGATTTTTCGGGCGCTCCACCGAA	240
Db	GGCTACGAATGATGGGCAAGCCCACTGTTTCCGACCTGATTTTTCGGGCGCTCCACCGAA	340
QY	ATGCTGCTCATCGCCCTCGTTGTGGGCGGAGCGCCCTGGGCGGCATCGCGCTCACACA	300
Db	ATGCTGCTCATCGCCCTCGTTGTGGGCGGAGCGCCCTGGGCGGCATCGCGCTCACACA	400
QY	TTGCTGTGTAATTCGCGCACGTAATCTAATGCGTTTCAATTCGCGCTGATGTGTCAA	360
Db	TTGCTGTGTAATTCGCGCACGTAATCTAATGCGTTTCAATTCGCGCTGATGTGTCAA	460
QY	AAACCCATTCGCCGTTTCTAATTCGGTTTCGGGCTTAATCGAAGCTACGAGTCACT	420
Db	AAACCCATTCGCCGTTTCTAATTCGGTTTCGGGCTTAATCGAAGCTACGAGTCACT	520
QY	GCGGCCAGGCGCGAGGCTGTGCGGCGTGGCACTATCTCAATCAATAATAGCGTTTCAC	480
Db	GCGGCCAGGCGCGAGGCTGTGCGGCGTGGCACTATCTCAATCAATAATAGCGTTTCAC	580
QY	TCTCACTGGGTAATTCGCGCGCTCTCACACGAGATGGCGATGCGAGATTATTCCTTGA	540
Db	TCTCACTGGGTAATTCGCGCGCTCTCACACGAGATGGCGAGATTAATTCCTTGA	640
QY	ATTAAAGGAGCTCGAGTTGCGCCTTTCCTCTCTTTGTGACAGCTACCTTGGATTCCGTC	600
Db	ATTAAAGGAGCTCGAGTTGCGCCTTTCCTCTCTTTGTGACAGCTACCTTGGATTCCGTC	700
QY	CGAACAAGAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGTTCAACATTTGCTCTT	660
Db	CGAACAAGAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGTTCAACATTTGCTCTT	760
QY	GTGGAATAATTCAGAGTCAAGGCCCTAATTGCGGCGCTGCTGATATTCTTGGGTCGTGTGAC	720
Db	GTGGAATAATTCAGAGTCAAGGCCCTAATTTGGGCGGCTGCTGATCTTCTTGGGTCGTGTGAC	820
QY	ATCCGTAATCTCTTCTTGGGAAAGGCTGTAA	753
Db	ATCCGTAATCTCTTCTTGGGAAAGGCTGTAA	853

QY 95 TCGGCAGAGGCTCAAAAACCTCCCTTGCTGTGAGATTGGGCGATTACCCGANTGGATATG 154
 Db 62 TCGCGAAGGCGTCAAAAGACAGTCTCCCATCGATCAGTACCTTCGGGTGGCATTCG 121
 QY 155 CGTTTGGTCTCTTGGTATTTCATCAACGCGTACGAATGTTGGGACGCCCACTGTTTTCG 214
 Db 122 CCTTCGGAGCTCAATGCTACCCGGGTGGCTTACCCCGCTGAAAGACTGTCTTCTCTCT 181
 QY 215 GCGTGAATTTGGCGGGGCTCCACCCGAATGCTGTGATCGCCCTGTTGTGGCGCGCAGCGC 274
 Db 182 GCATTATTTTAGCCGGGGCGGCGACGACATTGTGTGATCAGCGCATGTTGGCGCCCGGACAGTT 241
 QY 275 CCTTGGGCGCATGGCGCTCAACCAATTGCTGTGAACTTTCGGCCAGCTATTCTATGCGT 334
 Db 242 CGCGTGTGGGTGCGCGCCCTGACGGGTGATGAGGAGATGATGTCGCCACGATGCTGATGCGC 301
 QY 335 TTTCAATTC 343
 Db 302 CTTCACCTGC 310

APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

QY 241 ATGCTGATCGCCCTGTTGTGGGCGAGGCGCCCTGAGCCATCGCCTCAGACA 300
Db 241 ATGCTGATCGCCCTGTTGTGGGCGAGGCGCCCTGAGCCATCGCCTCAGACA 300
QY 301 TTGCTGTGAATCTTCGCGCAGTATCTATGCGTTTCAATCCGCTGATGTGTCAA 360
Db 301 TTGCTGTGAATCTTCGCGCAGTATCTATGCGTTTCAATCCGCTGATGTGTCAA 360
QY 361 AACCCATTTGCCGTTTCTATCGGTTTTCGCGTTTACGAGAGCCATCGAGTCACT 420
Db 361 AACCCATTTGCCGTTTCTATCGGTTTTCGCGTTTACGAGAGCCATCGAGTCACT 420
QY 421 GCGGCGAGGCGCGCAGGCTGTGTCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCA 480
Db 421 GCGGCGAGGCGCGCAGGCTGTGTCGCGTGGCGACTTATCTCAATGCAATAGCGTTTCA 480
QY 481 TCCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGAGTGAATTCCTTTGAA 540
Db 481 TCCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGAGTGAATTCCTTTGAA 540
QY 541 ATTAAGGCGCTCGAGTTCGCGCCCTTGTCTCTCTTGTGACGCTGACTTGGATTCTGCG 600
Db 541 ATTAAGGCGCTCGAGTTCGCGCCCTTGTCTCTCTTGTGACGCTGACTTGGATTCTGCG 600
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTGTCTT 660
Db 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTGTCTT 660
QY 661 GTGGTAATTCAGATCGAGGCGCTATTTGCGGCGCTGATCTCTTGGGTCTGTGAC 720
Db 661 GTGGTAATTCAGATCGAGGCGCTATTTGCGGCGCTGATCTCTTGGGTCTGTGAC 720
QY 721 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 753
Db 721 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 753

RESULT 2
US-09-471-803A-1
; Sequence 1, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGSELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; OTHER INFORMATION: ATCC14752
US-09-471-803A-1

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,9e-231;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATTCAGAGGCGCCCTG 60
Db 101 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATTCAGAGGCGCCCTG 160
QY 61 GAACCGATGATTAAGTTATTCGCGCTACGAAATCGCGCAAGTCTAAACCTCCCT 120
Db 161 GAACCGATGATTAAGTTATTCGCGCTACGAAATCGCGCAAGTCTAAACCTCCCT 220
QY 121 GCTGAGTTTGGGCAATACCCGATTTGATTTGCTTGTGTTATTCATAC 180
Db 221 GCTGAGTTTGGGCAATACCCGATTTGATTTGCTTGTGTTATTCATAC 280
QY 181 GGTACGATGTGGGAGCCCGACCTGTTTCCGCGCTGATTTTCCGCGGCTCCACCGAA 240
Db 281 GGTACGATGTGGGAGCCCGACCTGTTTCCGCGCTGATTTTCCGCGGCTCCACCGAA 340
QY 241 ATGCTGTGATCGCCCTGTTGTGGGCGAGCGCCCTGAGGCGCATCGGCTCAGACA 300
Db 341 ATGCTGTGATCGCCCTGTTGTGGGCGAGCGCCCTGAGGCGCATCGGCTCAGACA 400
QY 301 TTGCTGTGAATCTTCGCGCAGTATCTATGCGTTTCAATCCGCTGATGTGTCAA 360
Db 401 TTGCTGTGAATCTTCGCGCAGTATCTATGCGTTTCAATCCGCTGATGTGTCAA 460
QY 361 AACCCATTTGCCGTTTCTATTTGGTTTTCGCGCTTATGAGAGAGCTTACGAGTCACT 420
Db 461 AACCCATTTGCCGTTTCTATTTGGTTTTCGCGCTTATGAGAGAGCTTACGAGTCACT 520
QY 421 GCGGCGAGGCGCGCAGGCTGTGCGGCTGGCGACTTATCTCAATGCAATAGCGTTTCA 480
Db 521 GCGGCGAGGCGCGCAGGCTGTGCGGCTGGCGACTTATCTCAATGCAATAGCGTTTCA 580
QY 481 TCCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGAGTGAATTCCTTTGAA 540
Db 581 TCCCTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGAGTGAATTCCTTTGAA 640
QY 541 ATTAAGGCGCTCGAGTTCGCGCCCTTGTCTCTTGTGTCAGGCTGATTTGATTCCTG 600
Db 641 ATTAAGGCGCTCGAGTTCGCGCCCTTGTCTCTTGTGTCAGGCTGATTTGATTCCTG 700
QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTGTCT 660
Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTGTCT 760
QY 661 GTGGTAATTCAGATCGAGGCGCTATTTGCGGCGCTGATCTCTTGGGTCTGTGAC 720
Db 761 GTGGTAATTCAGATCGAGGCGCTATTTGCGGCGCTGATCTCTTGGGTCTGTGAC 820
QY 721 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 753
Db 821 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 853

RESULT 3
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGSELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1

Tue Apr 20 06:47:24 2004

us-10-608-504-6_copy_101_853.rn1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 63.4023 Seconds

(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753
Sequence: 1 gtgcagaaacgcagagat.....tctcggaagcgcctctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq:**
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq:**
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq:**
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq:**
- 5: /cgn2_6/prodata/2/ina/FCUS_COMB.seq:**
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	753	100.0	753	US-09-471-803A-2	Sequence 2, Appli
2	753	100.0	1271	US-09-471-803A-1	Sequence 1, Appli
3	753	100.0	1271	US-09-471-803A-6	Sequence 6, Appli
4	58.6	7.8	756	US-09-489-039A-3016	Sequence 3016, Ap
5	46.2	6.1	7218	US-08-232-463-14	Sequence 14, Appl
6	39.8	5.3	768	US-09-107-532A-1266	Sequence 1266, Ap
7	39.2	5.2	1038	US-09-252-991A-11376	Sequence 11376, A
8	39.2	5.2	1086	US-09-252-991A-11418	Sequence 11418, A
9	39.2	5.2	1399	US-09-621-976-8976	Sequence 8976, Ap
10	36.8	4.9	1140	US-09-352-991A-3205	Sequence 3205, Ap
11	36.8	4.9	2061	US-09-252-991A-2761	Sequence 2761, Ap
12	36.4	4.8	751	US-09-020-956-12	Sequence 12, Appl
13	36.4	4.8	751	US-09-030-607-12	Sequence 12, Appl
14	36.4	4.8	751	US-09-439-313-12	Sequence 12, Appl
15	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appl
16	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appl
17	36.4	4.8	751	US-09-159-812-12	Sequence 12, Appl
18	36.4	4.8	751	US-09-636-215-12	Sequence 12, Appl
19	36.4	4.8	751	US-09-685-166A-12	Sequence 12, Appl
20	36.4	4.8	751	US-09-115-453-12	Sequence 12, Appl
21	36.4	4.8	751	US-09-688-489-12	Sequence 12, Appl
22	35.4	4.7	430	US-09-621-976-16656	Sequence 16656, A
23	35.4	4.7	474	US-09-621-976-18033	Sequence 18033, A
24	35.2	4.7	801	US-09-328-352-2399	Sequence 2399, Ap
25	35.2	4.7	4403765	US-09-103-840A-2	Sequence 2, Appli
26	35.2	4.7	441529	US-09-103-840A-1	Sequence 1, Appli
27	34.8	4.6	867	US-09-489-039A-5047	Sequence 5047, Ap

28	34.6	4.6	666	2	US-08-875-034A-1	Sequence 1, Appli
29	34.4	4.6	744	4	US-09-489-039A-1307	Sequence 1307, Ap
30	34.4	4.6	1224	4	US-09-266-965-22	Sequence 22, Appl
31	34.4	4.6	12249	4	US-09-266-965-74	Sequence 74, Appl
32	34.4	4.6	18331	4	US-09-266-965-96	Sequence 96, Appl
33	33.4	4.4	1875	3	US-09-422-869-21	Sequence 21, Appl
34	33.4	4.4	2949	3	US-08-433-522A-1	Sequence 1, Appli
35	33.4	4.4	2949	3	US-09-135-166-1	Sequence 1, Appli
36	33.4	4.4	2949	3	US-08-942-046-1	Sequence 1, Appli
37	33.4	4.4	2950	3	US-08-433-522A-5	Sequence 5, Appli
38	33.4	4.4	2950	3	US-09-135-166-5	Sequence 5, Appli
39	33.4	4.4	2950	3	US-08-942-046-5	Sequence 5, Appli
40	33.4	4.4	2984	3	US-08-433-522A-3	Sequence 3, Appli
41	33.4	4.4	2984	3	US-09-135-166-3	Sequence 3, Appli
42	33.4	4.4	2984	3	US-08-942-046-3	Sequence 3, Appli
43	33.4	4.4	2987	3	US-08-433-522A-55	Sequence 55, Appli
44	33.4	4.4	2987	3	US-09-135-166-55	Sequence 55, Appli
45	33.4	4.4	2987	3	US-08-942-046-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1
US-09-471-803A-2
Sequence 2, Application US/09471803A
Parent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERNBECHT, NICOLE
APPLICANT: SAMM, HERMANN
APPLICANT: EGSELMING, LOTHAR
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471,803A
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentn Ver. 2.1
SEQ ID NO 2
LENGTH: 753
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
NAME/KEY: CDS
LOCATION: (1)..(753)
OTHER INFORMATION: DINF
OTHER INFORMATION: ATCC14752
US-09-471-803A-2

Query Match	100.0%;	Score 753;	DB 4;	Length 753;
Best Local Similarity	100.0%;	Pred. No. 2.2e+231;		
Matches 753;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCAAGCAGCCCTG	60	
Db	1	GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCAAGCAGCCCTG	60	
QY	61	GAACCAAGTATTAAGTTATCGGCGCTACGAATGCGGCAAGGTCTAANAACCTCCCTT	120	
Db	61	GAACCAAGTATTAAGTTATCGGCGCTACGAATGCGGCAAGGTCTAANAACCTCCCTT	120	
QY	121	GCTGCAAGTTGGCAGTATACCCGATTTGATTCGTTGCTCTTGTATTCAATAC	180	
Db	121	GCTGCAAGTTGGCAGTATACCCGATTTGATTCGTTGCTCTTGTATTCAATAC	180	
QY	181	GCTACGAATGTCGAGCCGACCCGCTTTTCGCGCTGATTTTCGCGGCTCCACCGAA	240	
Db	181	GCTACGAATGTCGAGCCGACCCGCTTTTCGCGCTGATTTTCGCGGCTCCACCGAA	240	

Db 18 YCWMSSGYMMWCTSGTY 2

RESULT 15

AAA10594/C
ID AAA10594 standard, DNA, 10732 BP.

XX AAA10594;

XX 29-JUN-2000 (first entry)

XX DE Gene encoding a subunit of cellulose synthase.

XX KM Cellulose synthase; cellulose production; increase yield; ds.

XX OS Vigna angularis.

XX PN JP2000060568-A.

XX PD 29-FEB-2000.

XX PF 26-AUG-1998; 98UP-00239998.

XX PR 26-AUG-1998; 98UP-00239998.

XX PA (MIZU) MIZUNO K.

XX PA (OUP) OUP PAPER CO.

XX DR WPI; 2000-342371/30.

XX P-PSDB; AAY85179.

XX PT A gene encoding a cellulose synthetic equipment - for the improvement in

XX PS the amount of cellulose synthesized in a plant body.

XX FS Claim 2, Page 14-21, 32pp; Japanese.

XX CC This sequence represents a gene encoding a subunit of the cellulose

XX CC synthase complex of Vigna angularis. The invention relates to subunits of

XX CC cellulose synthetic equipment, that can be used to increase the amount of

XX CC cellulose synthesized by a plant. The proteins and genes encoding them

XX CC can also be used to improve the properties of the cellulose being

XX CC produced by a plant

XX SQ Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 0 U; 2251 Other;

Query Match 5.5%; Score 41.2; DB 3; Length 10732;

Best Local Similarity 14.7%; Pred. No. 0.13;

Matches 64; Conservative 182; Mismatches 189; Indels 1; Gaps 1;

QY 262 GTGGGCGCAGCGCCCTGGGCGCCATCGCGCTCACCATTCGCTGTAAGTCCGCCAC 321

DB 9949 RYSSSTNCYSYCCYSRYSSTTNSRCCTTYSYSTBTBYSYSRGSRSRGS 9890

QY 322 GATATCTATGCGCTTTCATCCCGCTGATGTCGCAAAAACCCATGCCGCTTCTAT 381

DB 9889 RANCYNSNTNCYDASTDBYSRCCTTYSSTSTSTSTSTSTSTSTSTSTST 9830

QY 382 TCGGTTTGGCGCTTATCGACGAAGCCNACGATCAGTCCGCGCGCGCGCGCTGG 441

DB 9829 SDSTSTYRCRSRSDYDABSDNSTNCYDASRTBTBSTNCRCTBYDARCRSDSTY 9770

QY 442 TCGGCGTGGCGCTTATCGACGAAGCCNACGATCAGTCCGCGCGCGCGCTGG 501

DB 9769 GYDANSTYSRYSRYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSST 9710

QY 502 CTCACCGAGTGGCGATCGACAGATTGATCTTTGAAATPAAGGCGCTCG-AGTTGCG 560

DB 9709 YTBYSRRCRCYDAYSCRYDARCYDACSNSCTCYDATTSTSTSTSTSTST 9650

QY 561 CCTTGCTCTCTCTTGTGACGCTGACTTGATTCCTGCGGAGAAAGACAGATCCC 620

DB 9649 TTSRCTBT 9590

QY 621 TTCTCTGCTGCTGCGACGTTTGAGCTTCACATGCTCTTGNGTATTCAGGTGAGGC 680

DB 9589 CYSCTYDSDSTCYSTTYSSTYAYSCITTSRGSYDASTSTYBRTTSTYSSTYSDC 9530

QY 681 CCTATTGCGGCGCTG 696

DB 9529 SDYSTTBNTSTSSDS 9514

Search completed: April 19, 2004, 11:42:08
Job time : 352.669 secs

XX 05-OCT-1999; 99DE-01047792.
XX (DEGS) DEGUSA-HUELS AG.
XX Moeckel B, Pfeifferle W, Puenhler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX P-PsDB; AAB70881.
XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX Claim 4; Page 14-15; 22pp; German.
XX This invention describes a novel isolated nucleic acid (1) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (1) is used to
XX transform coryneforms for production of L-amino acids, specifically,
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (1) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX the Corynebacterium glutamicum lrp protein which is used in the method
XX described in the invention
XX Sequence 715 BP, 166 A, 173 C, 178 G, 198 T, 0 U, 0 Other;
SQ
XX
XX Query Match 11.3%; Score 85; DB 4; Length 715;
XX Best Local Similarity 100.0%; Pred. No. 9.6e-16;
XX Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTCAAAAAACGGAAGATTCATTCAAGCTCGAGGTCTCCCATCCAGGACACCTG 60
DB 85 GTTCAAAAAACGGAAGATTCATTCAAGCTCGAGGTCTCCCATCCAGGACACCTG 26
QY 61 GAACGAGATGATTAAGTTATCGGC 85
DB 25 GAACGAGATGATTAAGTTATCGGC 1
DB
XX
XX RESULT 14
XX ADA71938/c
XX ADA71938 standard; DNA; 2000 BP.
XX ADA71938;
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5263.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO200300898-A1.
XX
XX 03-JAN-2003.
XX
XX PD 22-JUN-2001; 2001WO-IB001105.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX

PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
SQ
XX
XX Query Match 6.2%; Score 47; DB 7; Length 2000;
XX Best Local Similarity 10.2%; Pred. No. 0.001;
XX Matches 75; Conservative 328; Mismatches 323; Indels 11; Gaps 2;
QY 8 AAACGAGAGATTCATTCAAGCTCGAGGTCTCCCATCCAGGACACCTGGAACG 67
DB 736 RMTMSAWRYCSRMKCATKYASAPWTAKRKYRFRYRMYRMYRKYRMYRMYRMYR 677
QY 68 ATGATTAAGTTATCGGCGGTACGAAATCCGGAAGCTTAAACCTCCTGCGAG 127
DB 676 RARMSKRRKWAAGSAMS CMYWRGASMYSKYSGACCKTKRWTSTSTSTSTSTST 617
QY 128 GTTTGGCATATGATCCGATGATGATGATGATGATGATGATGATGATGATGATG 187
DB 616 S--YKSMSTWSKMSYMGKMTCTMYTSKSTRSKMGMSGMYRMYRMYRMYRMYR 559
QY 188 AATGATGAGGACGCGGACCTGTTCCGCGCTGATTTCCGCGGCTCCACGGAATCTG 247
DB 558 YMKWKTWRRCMCYRMWYTYTTSRSRMYTGYKARYTSKRYMYRMYRMYRMYRMY 499
QY 248 TCATCGCCCT-----CGTTGGGCGCAGCGCCCTGCGCGCATCGCCTCACCA 298
DB 498 YMKCSYMMRYGYCAKCKCCYACWCWKAAYSMMWYRKYSKMYRMYRMYRMYRMY 439
QY 299 CATTGCTGTGAATTCGCGACGATATTCATGCGTTTATTCCGCGCTCATGTGTC 358
DB 438 SMKYGAGCGCGKMTYCSYGMKMYTMSYKYSRCYKMYRMYRMYRMYRMYRMY 379
QY 359 AAACCCGATGCGCGTTTCTATCGGTTTCCGCTTATCGAGCAAGCTTACGACTCA 418
DB 378 MMTYTYTAKMYKMYRGRGMSYKSKYKCTWMCYMKCRMYRMYRMYRMYRMYR 319
QY 419 CTGGCGCAGCGCGCAGGCTGTGCGCGTGGCGCATTCATCTCAATGCAATAGCTTC 478
DB 318 RCYCMRYATCYCCYKRGYVSRSRMTAGKMYRMSWRCGCSYMYRMYRMYRMY 259
QY 479 ACTGCTACTGGGTATTTGGGCGTCTCACCGAGTGGCGAGTGGAGTGAATCTCTTG 538
DB 258 YMSYGARSSGTRWSAARTRYKYSTSRARMMARARMTSACRMYRTSYCCGSCYGS 199
QY 539 AAATTAAGGCGCTGAGTGGCGCTTGTGCTCTCTTTGTCACGCTGACTTGGATTCT 598
DB 198 SKMYMSKSCMRMTCSMCSCYTCYGAACWMSCMMSMYMGSCGYTRGMKMRKRYSM 139
QY 599 GCGGACGAAAACAGATCCCTCTCTGCTGCTGCGCGAGTGGAGTTGAGCTTACATGCTC 658
DB 138 CCKKYCSCCTCYCSTGTGTYRKYKMYKYSYKCYCYCYWMSYRMYRMYRMYRMY 79
QY 659 TTGTGTATTCAGAGTCCGCTTATTTGGGCGCTGCTGATCTTGGGTGTGTA 718
DB 78 SCAYSTSTSTSRMMSMYAATGCGSSGMYRMSKMSKMSYSSCKYTGSKCKTKRY 19
QY 719 CCATCGGTAATCTTCTC 735

CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

CC
 XX Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 5; Length 309400;

Best Local Similarity 51.5%; Pred. No. 9.6e-28;

Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGACGATTGGGCGATGACCCGATGATGCTTGGTCTCTTGGTTTCAATACG 181
 DB 289969 CGGTGCGTTGGGGCTGATTCGCTGGGTTGGCTTTGGGCTGTGTGTCACACAG 289910

QY 182 GCTACGAATGATGGGAGAGCCGACGTTTCCGCGCTGATTTCCGCGGCTCAACCGAA 241

DB 289909 GTTTGCGCTGGTGGAGCGCGATTTCTCTGTGATCTATGCGGCTTCGATGAA 289850

QY 242 TCGTGTATCGCCCTGTTGGGCGAGCGCCCTGCGGCGCATCGGCTCAACCAT 301

DB 289849 TCTGGCAATCGGATGTCACCGAGGATCGGCCGTTTTCGCGCGGCTCGTGT 289790

QY 302 TCGTGTATCGCCGCGATGATCTATGCGTTTCAATCCGCGCTGATGTCGCAAA 361

DB 289789 TCAATGTGATTCGCGCAATTTCTAGGTCTCACTTCCACGACGCGCATCA 289730

QY 362 ACCCATTCGCGTTTCTATGCTTTCGCGCTTATGACGAGCTTACGAGCTACTG 421

DB 289729 CCGGCGCGCGCGCGCTATTCACCTACGCGCTTACGAGCAAGCTACGCGCAT 289670

QY 422 CGGCGAGGCGCGCGCTGTGGGCTG--GCGACTATCTCATGCAAAATAGGCTTC 478

DB 289669 CAGCGCGCGCGCGCTGATGATGAGCAGCGGAGTCTTACCGTCAAAATTTGTGCC 289610

QY 479 ACTCTACTGAGTATCGCGGCTTCAACCGAGTGGCGATCGCAGATGATTCCTTTTG 538

DB 289609 AAGCTGTGGGTATCCAGAAATATATGGGCGCTGTGTGTGAAGTGTGCCGATG 289550

QY 539 AAATTAAGGCGCTCGAGTTCGCGCTTGTCTCTCTTGTCAACGCTGACTTTGATTC 598

DB 289549 ATCTAAAGAGCATGATTTGTGCGCTGACGCGCTGTTGTGTGCTGGGAGGAG 289490

QY 599 GCGGAGCAAAAAGAGATCCCTTCTCTGCTGTCAGAGTTGAGCTTCAACATGCTC 658

DB 289489 TCAAAAATTAACAGATTTATGCTGCTCATATTCGCGGATGATGGCTGCTTCCG 289430

QY 659 TTGTGTAAATTCAGGTCAAGCCCTATTTGCGCGCTGCTGATCTTCTTGGTCTGTGA 718

DB 289429 GTTTGTGGCGCGCGAGAGATGCTGTATGCTTTGACACGATCTTTTGAATCTTC 289370

QY 719 CCATCGGTAATCTTCTC 735

DB 289369 TTCTCGCGCTCGCTTC 289353

RESULT 12

ACA01968
 ID ACA01968 standard; DNA; 177 BP.

ACA01968;

04-JUN-2003 (first entry)

C. glutamicum derived ORF SEQ ID 1959.

Coryneform; nucleic acid array; fermentation; culture; ds.

Corynebacterium glutamicum.

DE10128510-A1.

XX
 PD 19-DEC-2002.
 XX
 PF 13-JUN-2001; 2001DE-01028510.
 XX
 PR 13-JUN-2001; 2001DE-01028510.
 XX
 PA (DEBS) DEUS5A AG.
 XX
 PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;
 XX
 DR WPI; 2003-27970/28.
 XX

PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.

PS Claim 1; Page 647; 709pp; German.

CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA0001-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

SO Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;

Query Match 14.1%; Score 106; DB 7; Length 177;

Best Local Similarity 100.0%; Pred. No. 1.5e-22;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGAGTGTGCGATCCAGGCAAGCCCTG 60
 DB 72 GTGCAAAAACGCAAGATTCATTCAGGCTGAGAGTGTGCGATCCAGGCAAGCCCTG 131

QY 61 GAACCATGATTAAGGTATCGGCGCTACGAAATCGCGCAAGTTC 106

DB 132 GAACCATGATTAAGGTATCGGCGCTACGAAATCGCGCAAGTTC 177

RESULT 13

AA01968/C
 ID AA01968 standard; DNA; 715 BP.

AA01968;

12-JUN-2001 (first entry)

C. glutamicum lrp encoding DNA.

Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KW medicine; animal feed supplement; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers

-35_signal

-10_signal

CDS

EP1090993-A1.

11-APR-2001.

29-SEP-2000; 2000EP-00121159.

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;
SQ Query Match 73.4%; Score 553; DB 7; Length 564;
Best Local Similarity 99.8%; Pred. No. 2.7e-161;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 189 ATGTGGGAGGAGCCCACTGTTTTCGGGCTGATTTTCGGGAGCTCCACCGAAATGCTGCT 248
DB 1 ATGTGGGAGGAGCCCACTGTTTTCGGGCTGATTTTCGGGAGCTCCACCGAAATGCTGCT 60
QY 249 CATGCGCCCTGCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 308
DB 61 CATGCGCCCTGCTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 119
QY 309 GAACCTCCGCAAGATTTCTATGCGTTTTCATTCGCGCTGCAATGTTGTTCAAAAACCCCAT 368
DB 120 GAACCTCCGCAAGATTTCTATGCGTTTTCATTCGCGCTGCAATGTTGTTCAAAAACCCCAT 179
QY 369 TGGCGCTTCTATGCGTTTTCGCGCTGTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 428
DB 180 TGGCGCTTCTATGCGTTTTCGCGCTGTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 239
QY 429 GCGCGAGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 488
DB 240 GCGCGAGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 299
QY 489 GGTATTCGGCGGCTTCAACCGGAGTGGCGATTCGCAAGATTTCTGTTTGAATTAAGG 548
DB 300 GGTATTCGGCGGCTTCAACCGGAGTGGCGATTCGCAAGATTTCTGTTTGAATTAAGG 359
QY 549 CCTCGAGTTGCGCCTTGTGCTCTCTTTGTCAAGCTGACTTTGGAATTCCTGCGGAGCA 608
DB 360 CCTCGAGTTGCGCCTTGTGCTCTCTTTGTCAAGCTGACTTTGGAATTCCTGCGGAGCA 419
QY 609 AAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATTTGCTTGTAT 668
DB 420 AAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATTTGCTTGTAT 479
QY 669 TCCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATTTCTTGGTCTGTTGACCATCGGTA 728
DB 480 TCCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATTTCTTGGTCTGTTGACCATCGGTA 539
QY 729 CTTCCTTCTGGAAGAGGCTGCTAAA 753
DB 540 CTTCCTTCTGGAAGAGGCTGCTAAA 564

RESULT 7

AAF61693/c
ID AAF61693 standard; DNA; 397 BP.

XX AAF61693;

XX 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.

KM Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;

KW medicine; animal feed supplement; ds.

OS Corynebacterium glutamicum.

PN EPI090993-A1.

XX 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

PA (DEGUS) DEGUS-A-HUELS AG.

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI, 2001-29297/31.
PT New lrp gene from corynebacterium bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.

PS Example 3; Page 16; 22pp; German.

CC This invention describes a novel isolated nucleic acid (I) from
CC corynebacterium bacteria used for the fermentative production of selected L-
CC amino acids, by fermenting the amino acid-producing corynebacterium in which
CC at least the lrp gene has been weakened or amplified, then isolating
CC amino acids that have accumulated in the medium or cells. (II) is used to
CC transform corynebacterium for production of L-amino acids, specifically
CC lysine and isoleucine, which are used in medicine and particularly as
CC animal feed supplement. It may also be used as probes and primers for
CC isolating related sequences. Regulating expression of (I) improves
CC production of amino acids, especially of L-lysine. This sequence encodes
CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
CC the method described in the invention

SQ Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

Query Match 36.3%; Score 273.6; DB 4; Length 397;
Best Local Similarity 98.6%; Pred. No. 1.7e-74;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAAGATTTCAATTCAGGCTGAGAGTGGCGGCTGCAAGGAGGAGGAGGAG 60
DB 284 GTGCAAAAAGCAAGATTTCAATTCAGGCTGAGAGTGGCGGCTGCAAGGAGGAGGAGGAG 225
QY 61 GAACCAAGATTAAGGTTATTCGCGCTGCAAGATTCGCGGCTGCAAGGAGGAGGAGGAG 120
DB 224 GAACCAAGATTAAGGTTATTCGCGCTGCAAGATTCGCGGCTGCAAGGAGGAGGAGGAG 165
QY 121 GCTGAGGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 164 GCTGAGGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 105
QY 181 GGTACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
DB 104 GGTACCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45
QY 241 ATGCTGCTATCGCCCTGCTGTTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 280
DB 44 ATGCTGCTATCGCCCTGCTGTTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5

RESULT 8

AAF61695/c
ID AAF61695 standard; DNA; 778 BP.

XX AAF61695;

XX 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.

KM Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;

KW medicine; animal feed supplement; ds.

OS Corynebacterium glutamicum.

PN EPI090993-A1.

XX 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

PA (DEGUS) DEGUS-A-HUELS AG.

XX C glutamicum coding sequence fragment SEQ ID NO: 1.
 XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 XX Corynebacterium glutamicum.
 PN Epi108790-A2.
 PD 20-JUN-2001.
 XX 18-DEC-2000; 2000EP-00127688.
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX (KYOM) KYOMA HAKKO KOGYO KK.
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda W, Ozaki A;
 DR WPI; 2001-376931/40.
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 PS Claim 7, SEQ ID NO 1; 246pp + Sequence listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
 CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 CC
 SQ Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;
 Query Match 100.0%; Score 753; DB 5; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 3,1e-222;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGAGATTCAATTCAGCCTGAGAGTGTGCGCATCCAGGACGCTCG 60
 DB 276829 GTGCAAAAACGCAAGAGATTCAATTCAGCCTGAGAGTGTGCGCATCCAGGACGCTCG 276888
 QY 61 GAACGAGATGATTAAGGTATCGGCGCTACGAATCGGCGAGGCTTAAAACTCCCTT 120
 DB 276889 GAACGAGATGATTAAGGTATCGGCGCTACGAATCGGCGAGGCTTAAAACTCCCTT 276948
 QY 121 GCTGAGGTTTGGGCGATGATCCGATGATGCGTTTGTCTCTTGGTTATTCATAC 180
 DB 276949 GCTGAGGTTTGGGCGATGATCCGATGATGCGTTTGTCTCTTGGTTATTCATAC 277008
 QY 181 GGTAGAGATGATGAGGCGACGCCCATGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
 DB 277009 GGTAGAGATGATGAGGCGACGCCCATGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 277068
 QY 241 ATGCTGTATGATCGCTCGTTTGGGCGAGCGCCCTTGGGCGCATCGCGCTCACCA 300
 DB 277069 ATGCTGTATGATCGCTCGTTTGGGCGAGCGCCCTTGGGCGCATCGCGCTCACCA 277128
 QY 301 TTGCTGTATGATCGCGACGATTCATGCTTTCATCCCGCTGATGATGATCAAA 360
 DB 277129 TTGCTGTATGATCGCGACGATTCATGCTTTCATCCCGCTGATGATGATCAAA 277188

QY 361 AACCCATGTCCTGTTTCTATTCGCTTTGCGCTTATGACGAAAGCTTACGAGTCACT 420
 DB 277189 AACCCATGTCCTGTTTCTATTCGCTTTTTCGCTTATGACGAAAGCTTACGAGTCACT 277248
 QY 421 GCGGCGAGGCGCGGCGGCTGCTCGGCGTGGGCGACTTATCTCAATGCAATAGGCTTTCAC 480
 DB 277249 GCGGCGAGGCGCGGCGGCTGCTCGGCGTGGGCGACTTATCTCAATGCAATAGGCTTTCAC 277308
 QY 481 TCCTACTGAGTATTCGCGGCTGCTCACCGGAGTGGCGATGCAAGATTCTTTTGA 540
 DB 277309 TCCTACTGAGTATTCGCGGCTGCTCACCGGAGTGGCGATGCAAGATTCTTTTGA 277368
 QY 541 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTCTTGTGACGCTGACTTGAATTCCTG 600
 DB 277369 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTCTTGTGACGCTGACTTGAATTCCTG 277428
 QY 601 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTCAACATTGCTCTT 660
 DB 277429 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTCAACATTGCTCTT 277488
 QY 661 GTGTAATTCAGATGAGGCTTATTTGGGCGGCTGATCTTCTTGGTCTGTTGAC 720
 DB 277489 GTGTAATTCAGATGAGGCTTATTTGGGCGGCTGATCTTCTTGGTCTGTTGAC 277548
 QY 721 ATCCGACTTCTTCTTGGGAAAGGCTGCTAAA 753
 DB 277549 ATCCGACTTCTTCTTGGGAAAGGCTGCTAAA 277581
 RESULT 6
 ID ACA01969 standard; DNA; 564 BP.
 XX ACA01969;
 AC 04-JUN-2003 (first entry)
 DT 04-JUN-2003 (first entry)
 XX C. glutamicum derived ORF SEQ ID 1960.
 DE C. glutamicum derived ORF SEQ ID 1960.
 KM Coryneform; nucleic acid array; fermentation; culture; ds.
 XX Corynebacterium glutamicum.
 OS Corynebacterium glutamicum.
 PN DE10128510-A1.
 FN 19-DEC-2002.
 PD 13-JUN-2001; 2001DE-01028510.
 PF 13-JUN-2001; 2001DE-01028510.
 XX 13-JUN-2001; 2001DE-01028510.
 FR (DEGS) DEGUSA AG.
 PA Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
 XX WPI; 2003-279970/28.
 DR WPI; 2003-279970/28.
 XX New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 PS Claim 1; Page 647; 709pp; German.
 XX This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA0010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

Db 641 ATTAAGGAGGCTGAGATTCGCTTGTCTCTCTTTCACGCTGATTCGATTCCTGC 700
Qy 601 CGAAGAAAAGAGATCCCTTCTGCTGCTCGAGATTGACCTTACCTTCTT 660
Db 701 CGAAGAAAAGAGATCCCTTCTGCTGCTCGAGATTGACCTTACCTTCTT 760
Qy 661 GTGTAATTCAGGTCAGGCTTATTCGAGGCTGATCTCTTGGCTGTGACC 720
Db 761 GTGTAATTCAGGTCAGGCTTATTCGAGGCTGATCTCTTGGCTGTGACC 820
Qy 721 ATCCGTAATTCCTTCTTGGGAAAAGGCTGTAAA 753
Db 821 ATCCGTAATTCCTTCTTGGGAAAAGGCTGTAAA 853

RESULT 4

AAH21112 standard; DNA; 1271 BP.

AAH21112;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnB.

L-amino acid production; brnF, brnB; branched-chain amino acid;
coryneform bacterium; leucine; isoleucine; valine; medicine;
animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers

CDS 101..856

FT /*tag= a

FT /product= "brnF"

FT 853..1179

FT /*tag= b

FT /product= "brnB"

EP1096010-A1.

02-MAY-2001.

11-OCT-2000; 2000BP-00122057.

27-OCT-1999; 99DE-01051708.

(DEGS) DEGUSSA AG.

(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

Kernermecht N. Eggeling L. Sahm H. Pfefferle W.

MPI; 2001-391595/42.

New export genes from coryneform bacteria, useful for increasing

fermentative production of branched-chain amino acids.

Claim 4 (i); Page 17-18; 23pp; German.

This invention describes a novel isolated polynucleotide (I) containing
at least one sequence that (i) is 70% identical with a sequence that
encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
15 consecutive bases from (i)-(iii). The invention also describes (a)
protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
coryneform microorganisms, especially Corynebacterium, transformed with
one or more (i), where these are replicative DNA; (c) production of
branched-chain L-aa by fermentation of coryneform bacteria in which the
brnB and/or brnF genes (or equivalent sequences) are amplified,
especially overexpressed; and (d) method for isolating the brnB and/or
brnF genes. (i) is used for transformation of coryneform bacteria being
used for fermentative production of branched-chain amino acids.

CC Specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (i) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (i)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 13032 brnF and brnB proteins described in
CC the method of the invention
XX
SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,55-223; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

Qy 1 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATCCAGGCAAGCCCTG 60
Db 101 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATCCAGGCAAGCCCTG 160
Qy 61 GAACCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 120
Db 161 GAACCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 220
Qy 121 GCTGCAAGTTTGGGCAATGACCGATTTGATTTGATTTGATTTGATTTGATTTG 180
Db 221 GCTGCAAGTTTGGGCAATGACCGATTTGATTTGATTTGATTTGATTTGATTTG 280
Qy 181 GGCTACGAATGTGTGGGCAAGCCCACTTTTTCGGGCTGATTTTTCGGGCTGATTT 240
Db 281 GGCTACGAATGTGTGGGCAAGCCCACTTTTTCGGGCTGATTTTTCGGGCTGATTT 340
Qy 241 ATGCTGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 341 ATGCTGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
Qy 301 TTGCTGATTAAGATTCGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 360
Db 401 TTGCTGATTAAGATTCGCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 460
Qy 361 AACCCCAATGCGGCTTCTATTCATTCATTCATTCATTCATTCATTCATTCATTC 420
Db 461 AACCCCAATGCGGCTTCTATTCATTCATTCATTCATTCATTCATTCATTCATTC 520
Qy 421 GCGGCCAGGCGCGAGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCG 480
Db 521 GCGGCCAGGCGCGAGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCG 580
Qy 481 TCCACTGAGGATTCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTG 540
Db 581 TCCACTGAGGATTCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTG 640
Qy 541 ATTAAGGAGCTGAGATTCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 600
Db 641 ATTAAGGAGCTGAGATTCGCGGCTGTGCGGCTGTGCGGCTGTGCGGCTGTGCGG 700
Qy 601 CGAAGCAAAAAGAGATCCCTTCTGCTGCTCGAGATTGACCTTACCTTCTT 660
Db 701 CGAAGCAAAAAGAGATCCCTTCTGCTGCTCGAGATTGACCTTACCTTCTT 760
Qy 661 GTGTAATTCAGGTCAGGCTTATTCGAGGCTGATCTCTTGGCTGTGACC 720
Db 761 GTGTAATTCAGGTCAGGCTTATTCGAGGCTGATCTCTTGGCTGTGACC 820
Qy 721 ATCCGTAATTCCTTCTTGGGAAAAGGCTGTAAA 753
Db 821 ATCCGTAATTCCTTCTTGGGAAAAGGCTGTAAA 853

RESULT 5

AAH64966 standard; DNA; 349980 BP.

AAH64966;

26-SEP-2001 (first entry)

```

DB 241 ATGCTGATACGCTCTGTTGAGGCGACGCGCCCTGGGCGGCATCGGCTCACCA 300
QY 301 TTGCTGATGAACTTCGCGCAAGATATCTATGCGTTTCATTCGCGTGCATGTGTCAA 360
DB 301 TTGCTGATGAACTTCGCGCAAGATATCTATGCGTTTCATTCGCGTGCATGTGTCAA 360
QY 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCG 420
DB 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCG 420
QY 421 GCGGCGACGCGCGCGAGCTGCGCGCGTGCAGTATCTCAATGCAATAGCGTTTCA 480
DB 421 GCGGCGACGCGCGCGAGCTGCGCGCGTGCAGTATCTCAATGCAATAGCGTTTCA 480
QY 481 TCCTACTGAGTATTCGCGCGTCTCAACGAGATGCGAGTGCAGATGAGTTGATTCCTTGA 540
DB 481 TCCTACTGAGTATTCGCGCGTCTCAACGAGATGCGAGTGCAGATGAGTTGATTCCTTGA 540
QY 541 ATTAAGGCGCTGAGTTCGCGCTTTCCTCTCTTTTTCAGCGTGCATTTGATTCCTTGC 600
DB 541 ATTAAGGCGCTGAGTTCGCGCTTTCCTCTCTTTTTCAGCGTGCATTTGATTCCTTGC 600
QY 601 CGAAGAAAAGAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTT 660
DB 601 CGAAGAAAAGAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTT 660
QY 661 GTGGTAATTCAGAGTGCAGCGCCCTATTTGCGGCGGCTGATCTTCTGCGTCTTGA 720
DB 661 GTGGTAATTCAGAGTGCAGCGCCCTATTTGCGGCGGCTGATCTTCTGCGTCTTGA 720
QY 721 ATCCGTAATTCCTTCTTGGGAAAGCGCTGCTAA 753
DB 721 ATCCGTAATTCCTTCTTGGGAAAGCGCTGCTAA 753
RESULT 3
AAH21109 standard; DNA; 1271 BP.
AAH21109;
05-SEP-2001 (first entry)
C. glutamicum DNA encoding brnF and brnE.
L-amino acid production; brnF; brnE; branched-chain amino acid;
corynebacterium; leucine; isoleucine; valine; medicine;
animal nutrition; ds.
Corynebacterium glutamicum.
Key Location/Qualifiers
CDS 101..856
FT /tag= a
FT /product= "brnF"
FT 853..1179
FT /tag= b
FT /product= "brnE"
EP1096010-A1.
02-MAY-2001.
11-OCT-2000; 2000BP-00122057.
27-OCT-1999; 99DE-01051708.
PA (DEGS) DEGUSA AG.
PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX Kennernknecht N, Eggeling L, Sahm H, Pfeifferle W;
XX

```

```

DR WPI; 2001-391595/42.
DR P-PSDB; AAB86247, AAB86248.
PT New export genes from corynebacterium bacteria, useful for increasing
PT fermentative production of branched-chain amino acids.
XX Claim 4 (1); Page 13; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (1) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC corynebacterium microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of
CC branched-chain L-aa by fermentation of corynebacterium bacteria in which the
CC brnE and/or brnF genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the brnE and/or
CC brnF genes. (1) is used for transformation of corynebacterium bacteria being
CC used for fermentative production of branched-chain amino acids;
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC corynebacterium glutamicum ATCC 14752 brnF and brnE proteins described in
CC the method of the invention
XX
SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;
Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,5e-22; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;
QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCATCAAGGAGCGCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCATCAAGGAGCGCTG 160
QY 61 GAAACAGATGATTAAGATTATCGGCGCTACGAAATGCGCAAGCTTAAGACCTCCCT 120
DB 161 GAAACAGATGATTAAGATTATCGGCGCTACGAAATGCGCAAGCTTAAGACCTCCCT 220
QY 121 GCTGCAAGTTTGGGCAATGATACCGGATTTGATTCGCTTCTTGTATTCATAC 180
DB 221 GCTGCAAGTTTGGGCAATGATACCGGATTTGATTCGCTTCTTGTATTCATAC 280
QY 181 GGCTACGAATGATGAGGAGCGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
DB 281 GGCTACGAATGATGAGGAGCGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 340
QY 241 ATGCTGATTCAGCGCCCTGTTGAGGAGCGCGCCCTGAGGCGCATTCGCGCTCACCA 300
DB 341 ATGCTGATTCAGCGCCCTGTTGAGGAGCGCGCCCTGAGGCGCATTCGCGCTCACCA 400
QY 301 TTGCTGATGAACTTCGCGCAAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAA 360
DB 401 TTGCTGATGAACTTCGCGCAAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAA 460
QY 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCG 420
DB 461 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCG 520
QY 421 GCGGCGACGCGCGAGCTGCGCGCGTGCAGTATCTCAATGCAATAGCGTTTCA 480
DB 521 GCGGCGACGCGCGAGCTGCGCGCGTGCAGTATCTCAATGCAATAGCGTTTCA 580
QY 481 TCCTACTGAGTATTCGCGCGTCTCAACGAGATGCGAGTGCAGATGAGTTGATTCCTTGA 540
DB 581 TCCTACTGAGTATTCGCGCGTCTCAACGAGATGCGAGTGCAGATGAGTTGATTCCTTGA 640
QY 541 ATTAAGGCGCTGAGTTCGCGCTTTCCTCTCTTTTTCAGCGTGCATTTGATTCCTTGC 600

```

branched-chain L-aa by fermentation of coryneform bacteria in which the
CC brn and/or brn genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the brn and/or
CC brn genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 14752 brnF protein described in the
CC method of the invention

Sequence 753 BP, 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 753;

Best Local Similarity 100.0%; Pred. No. 2e-223; Mismatches 0; Indels 0; Gaps 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGCAAGATTCATTCAAGCTGAGAGTGGCCATCCAAAGCAGCCCTG 60
1 GTGCAAAAACCGCAAGATTCATTCAAGCTGAGAGTGGCCATCCAAAGCAGCCCTG 60
61 GAACGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGTTTAAACCTCCCTT 120
61 GAACGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGTTTAAACCTCCCTT 120
121 GCTGCAGTTTGGGCAATGACCGGATGTTGCTTGGCTTGGTTATTCATATC 180
121 GCTGCAGTTTGGGCAATGACCGGATGTTGCTTGGCTTGGTTATTCATATC 180
181 GGCTACGAATGATGAGGAGCCCACTGTTTCCGGCTGATTTTCCGGCTGACCGAA 240
181 GGCTACGAATGATGAGGAGCCCACTGTTTCCGGCTGATTTTCCGGCTGACCGAA 240
241 ATGCTGATCATCGGCTGTTGTTGGGCGCAGGCGCCCTGGGCGCATCCGCTCACCA 300
241 ATGCTGATCATCGGCTGTTGTTGGGCGCAGGCGCCCTGGGCGCATCCGCTCACCA 300
301 TTGCTGTAAGTTCGCGCAGATTCATGCGTTTCAATCCCGCTGATGTTGTTCA 360
301 TTGCTGTAAGTTCGCGCAGATTCATGCGTTTCAATCCCGCTGATGTTGTTCA 360
361 AACCCATTCGCGCTTCTATTCGCTTTCGCGTTATGCAAGAACCTTACGATGACT 420
361 AACCCATTCGCGCTTCTATTCGCTTTCGCGTTATGCAAGAACCTTACGATGACT 420
421 GCGGCGAGGCGCGAGGCTGGTGGCGGCTGCGGATTCATCAATGCAAAAGGTTTAC 480
421 GCGGCGAGGCGCGAGGCTGGTGGCGGCTGCGGATTCATCAATGCAAAAGGTTTAC 480
481 TCCTAGTGGATTCGCGGCTTCACCGAGTGGCGATTCGAGATTCGATTCCTTTGA 540
481 TCCTAGTGGATTCGCGGCTTCACCGAGTGGCGATTCGAGATTCGATTCCTTTGA 540
541 ATTAAGGCTCGAGTTCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 600
541 ATTAAGGCTCGAGTTCGCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 600
601 CGAAGCAAAAACGAGATTCCTTCTCTGCTGCTCGAGGTTTGAAGCTTACCATTTG 660
601 CGAAGCAAAAACGAGATTCCTTCTCTGCTGCTCGAGGTTTGAAGCTTACCATTTG 660
661 GTGTAATTCGAGGAGGCGCTTTCGCGGCTGCGGATTCCTTGGGCTGTTGAC 720
661 GTGTAATTCGAGGAGGCGCTTTCGCGGCTGCGGATTCCTTGGGCTGTTGAC 720
721 ATCCGTAATTCCTTCTTGGGAAAGGCTGCTAAA 753
721 ATCCGTAATTCCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 2
AAH65254

ID AAH65254 strand; DNA; 753 BP.

XX AAH65254;

XX 26-SEP-2001 (first entry)

DT C glutamicum coding sequence fragment SEQ ID NO: 289.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KJOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI: 2001-376931/40.

XX P-FSDB; AAG90035.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 8; SEQ ID NO 289; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

XX sequences from the Corynebacterium glutamicum. These

XX are useful for identifying the mutation point of a gene derived from a

XX mutant of corynebacterium, measuring expression amount and analyzing

XX the expression profile or expression pattern of a gene derived from

XX corynebacterium, and identifying a homologue of a gene derived from

XX corynebacterium. Corynebacterium bacteria are useful for producing amino

XX acids, nucleic acids, vitamins, saccharides and organic acids,

XX particularly L-lysine. The present sequence is a nucleic acid described

XX in the exemplification of the invention. Note: The sequence data for this

XX patent did not form part of the printed specification, but was obtained

XX in electronic format directly from the European Patent Office

XX Sequence 753 BP, 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 753; DB 5; Length 753;

XX Best Local Similarity 100.0%; Pred. No. 2e-223; Mismatches 0; Indels 0; Gaps 0;

XX Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGCAAGATTCATTCAAGCTGAGAGTGGCCATCCAAAGCAGCCCTG 60
1 GTGCAAAAACCGCAAGATTCATTCAAGCTGAGAGTGGCCATCCAAAGCAGCCCTG 60
61 GAACGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGTTTAAACCTCCCTT 120
61 GAACGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGTTTAAACCTCCCTT 120
121 GCTGCAGTTTGGGCAATGACCGGATGTTGCTTGGCTTGGTTATTCATATC 180
121 GCTGCAGTTTGGGCAATGACCGGATGTTGCTTGGCTTGGTTATTCATATC 180
181 GGCTACGAATGATGAGGAGCCCACTGTTTCCGGCTGATTTTCCGGCTGACCGAA 240
181 GGCTACGAATGATGAGGAGCCCACTGTTTCCGGCTGATTTTCCGGCTGACCGAA 240
241 ATGCTGATCATCGGCTGTTGTTGGGCGCAGGCGCCCTGGGCGCATCCGCTCACCA 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 350.669 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853
Perfect score: 753
Sequence: 1 gtgcacaaacgcgaagat.....tcttgggaagcctctaaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29yand4:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	4 AAH21110	Aah21110 C. glutam
2	753	100.0	753	5 AAH65254	Aah65254 C. glutam
3	753	100.0	1271	4 AAH21109	Aah21109 C. glutam
4	753	100.0	1271	4 AAH21112	Aah21112 C. glutam
5	753	100.0	349860	4 AAH64966	Aah64966 C. glutam
6	553	73.4	564	7 ACA01969	ACA01969 C. glutam
7	273.6	36.3	397	4 AAH61693	AAH61693 C. glutam
8	273.6	36.3	778	4 AAH61695	AAH61695 C. glutam
9	127.4	16.9	711	5 AAH68373	AAH68373 C. glutam
10	127.4	16.9	711	5 AAH68373	AAH68373 C. glutam
11	127.4	16.9	309460	5 AAH68534	Aah68534 C. glutam
12	106	14.1	177	7 ACA01968	ACA01968 C. glutam
13	85	11.3	715	4 AAH61688	AAH61688 C. glutam
14	47	6.2	2000	7 ADA71938	Ada71938 Rice gene
15	41.2	5.5	10732	3 AAA10594	AAA10594 Gene enco
16	40.8	5.4	7645	9 ADB59610	ADB59610 Human gen
17	40.2	5.3	4947	2 ADB59610	ADB59610 Human gen
18	39.8	5.3	768	9 ADB59610	ADB59610 Human gen
19	37.6	5.0	759	9 ADB59610	ADB59610 Human gen
20	37.6	5.0	5059	2 AAH44332	AAH44332 Streptococcus
21	37.6	5.0	11000	7 ACFe7367_20	ACFe7367_20 Continuation (21 o
22	37.6	5.0	243072	7 ACFe7367_20	ACFe7367_20 Continuation (21 o
23	37.2	4.9	1024	6 ABX66366	ABX66366 Helicobac

24	36.8	4.9	1230	7 ACA42205	ACA42205 Prokaryot
25	36.4	4.8	751	2 AAH58489	AAH58489 3' fragme
26	36.4	4.8	751	2 AAH58489	AAH58489 3' fragme
27	36.4	4.8	751	3 AAA06252	AAA06252 Human imm
28	36.4	4.8	751	3 AAH71157	AAH71157 Human pro
29	36.4	4.8	751	3 AAH93368	AAH93368 Human pro
30	36.4	4.8	751	4 AAH63460	AAH63460 Human pro
31	36.4	4.8	751	4 AAH02433	AAH02433 Prostate
32	36.4	4.8	751	4 AAH04682	AAH04682 Human pro
33	36.4	4.8	751	5 ACA59269	ACA59269 Prostate
34	36.4	4.8	751	5 AAH10011	AAH10011 Human pro
35	36.4	4.8	751	6 ABH4832	ABH4832 Human N1-
36	36.4	4.8	751	6 ABH58541	ABH58541 Prostate
37	36.4	4.8	751	7 ACC94996	ACC94996 Prostate
38	36.4	4.8	751	9 ADB13462	ADB13462 Human pro
39	35.6	4.7	13638	7 ABX34289	ABX34289 S. atrool
40	35.4	4.7	1362	7 ADA70464	Ada70464 Rice gene
41	35.2	4.7	693	7 ABZ38247	ABZ38247 N. gonorr
42	35.2	4.7	801	8 ADA31112	ADA31112 DNA encod
43	35.2	4.7	2001	4 AAH61084	AAH61084 P. putida
44	35.2	4.7	2028	4 AAH61083	AAH61083 P. putida
45	35.2	4.7	11000	4 AAH9682_33	AAH9682_33 Continuation (34 o

ALIGNMENTS

RESULT 1	AAH21110	standard; DNA; 753 BP.
AC	AAH21110;	
DT	05-SEP-2001	(first entry)
DE	C. glutamicum bnf DNA.	
KE	L-amino acid production; bnf, bnf, branched-chain amino acid;	
KW	coryneform bacterium; leucine; isoleucine; valine; medicine;	
KW	animal nutrition; ds.	
OS	Corynebacterium glutamicum.	
PN	EP1096010-A1.	
XX	02-MAY-2001.	
PF	11-OCT-2000; 2000EP-00122057.	
XX	27-OCT-1999; 99DE-01051708.	
PA	(DEGS) DEGUSA AG.	
PA	(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.	
PI	Kennernrecht N, Eggeling L, Sahm H, Pfeifferle W;	
DR	WPI; 2001-391595/42.	
DR	P-Psdb; AAB86247.	
PT	New export genes from coryneform bacteria, useful for increasing	
PT	fermentative production of branched-chain amino acids.	
XX	Claim 5; Page 14-15; 23pp; German.	
PS	This invention describes a novel isolated polynucleotide (1) containing	
XX	at least one sequence that (i) is 70% identical with a sequence that	
CC	encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)	
CC	polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)	
CC	or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least	
CC	15 consecutive bases from (i)-(iii). The invention also describes (a)	
CC	protein derived from sequences of 753 bp (2) or 324 bp (4); (b)	
CC	coryneform microorganisms, especially corynebacterium, transformed with	
CC	one or more (i), where these are replicative DNA; (c) production of	

Db 277489 GTGGTAATTCAGGTCAGGCTATTTCGGCGCTGATCTTCTGGTCTGTGACC 277548
QY 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 753
Db 277549 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 277581

RESULT 15

AX137083/c 397 bp DNA Linear PAT 30-MAY-2001
LOCUS AX137083
DEFINITION Sequence 7 from Patent EP1090993.
ACCESSION AX137083
VERSION AX137083.1 GI:14273428
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
1 Moeckel, B., Pfeifferle, W., Puenler, A., Kalinowski, J. and Bathe, B.
TITLE Nucleotide sequences coding for the lrp gene
JOURNAL Patent: EP 1090993-A 7 11-APR-2001;
Degussa AG (DE)

FEATURES
Source Location/Qualifiers
1..397
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="lrp-Teil 1"

ORIGIN

Query Match 36.3%; Score 273.6; DB 6; Length 397;
Best Local Similarity 98.6%; Pred. No. 3.8e-60;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTTCATTCAGCTGAGGTGTGCGCATCCAGGAGCCCTG 60
284 GTGCAAAAAACGCAAGATTTCATTCAGCTGAGGTGTGCGCATCCAGGAGCCCTG 225
QY 61 GAACCAATGATTAAGTTATGCGCGCTACGAAATGCGCAAGGCTTAAACCTCCCTT 120
Db 224 GAACCAATGATTAAGTTATGCGCGCTACGAAATGCGCAAGGCTTAAACCTCCCTT 165
QY 121 GTGCAAGTTTGGCATGATCCGATGATGCGTTGCTCTCTTGGTATTCATATAC 180
Db 164 GCTGCAAGTTTGGCATGATCCGATGATGCGTTGCTCTCTTGGTATTCATATAC 105
QY 181 GGTACCAATGATGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
Db 104 GGTACCAATGATGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 45
QY 241 ATGCTGTCATCGCCCTCGTTGTGAGGAGCAGCCCTCTG 280
Db 44 ATGCTGTCATCGCCCTCGTTGTGAGGAGCAGCCCTCTG 5

Search completed: April 19, 2004, 18:30:10
Job time : 3194.51 secs


```

/translation="MTSLRDLVSTNADAIADMSAFIEETVKGSGIGATKANG
AATKIDSDIVAKGSRRLIPRIADSLDGMQRYNGCTAGLHLENNSSPALDAILIS
VADKNABESINVEGLKATYKYGKRAAKYIEDELPIGLIERNKQ"
gene
complement(10104..11177)
/gene="Cg10010"
complement(10104..11177)
/gene="Cg10010"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAB97403.1"
/db_xref="GI:21322774"
/translation="MFMLAORTLPPIHTAPHLFVARVPHQIRATDADRTSLORDLELS

Query Match      100.0%; Score 753; DB 1; Length 340000;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGCAGCCCTG 60
Db 276829 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGCAGCCCTG 276888
QY 61 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 120
Db 276889 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 276948
QY 121 GCTGCAAGTTGGGCAATGACCCGATGATGCTTGGTGTGCTCTGCTTGGTTATTCATATAC 180
Db 276949 GCTGCAAGTTGGGCAATGACCCGATGATGCTTGGTGTGCTCTGCTTGGTTATTCATATAC 277008
QY 181 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTGATTTTGGCGGGCTCCACCGAA 240
Db 277009 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTGATTTTGGCGGGCTCCACCGAA 277068
QY 241 ATGCTGATCATGCGCCCTGTTGAGGCGGCAAGCCCTGAGGCGCATGCGCTCACCA 300
Db 277069 ATGCTGATCATGCGCCCTGTTGAGGCGGCAAGCCCTGAGGCGCATGCGCTCACCA 277128
QY 301 TTGCTGTGAATCTTCCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAA 360
Db 277129 TTGCTGTGAATCTTCCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAA 277188
QY 361 AACCCCATGCGCGTTTCTATTCGCTTTCGCGCTTATGAGAGCCATGACAGTCACT 420
Db 277189 AACCCCATGCGCGTTTCTATTCGCTTTCGCGCTTATGAGAGCCATGACAGTCACT 277248
QY 421 GCGGCAAGGCGCGAGGCTGTGCGGCTGCGGCTTATGCTCAATGCAATAGCGTTTAC 480
Db 277249 GCGGCAAGGCGCGAGGCTGTGCGGCTGCGGCTTATGCTCAATGCAATAGCGTTTAC 277308
QY 481 TCCACTGAGGTATTCGCGGCTGCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 540
Db 277309 TCCACTGAGGTATTCGCGGCTGCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 277368
QY 541 ATTAAGGCGCTCGAGTTTCGCGCTTTCGCTTTCGCTTTCGCTGACGCTGATTCCTGC 600
Db 277369 ATTAAGGCGCTCGAGTTTCGCGCTTTCGCTTTCGCTTTCGCTGACGCTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 660
Db 277429 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 277488
QY 661 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTTCTTGGGTCTGTGACC 720
Db 277489 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTTCTTGGGTCTGTGACC 277548
QY 721 ATCCGTAATCTTCTTCTTGGGAAAAGGCTGCTAAA 753
Db 277549 ATCCGTAATCTTCTTCTTGGGAAAAGGCTGCTAAA 277581

```

```

LOCUS AX120085 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent EP1108790.
ACCESSION AX120085 AX114121
VERSION AX120085.1 GI:114036800
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source
1..349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/notes="Seq 1 to long (3,309,400) split in 11, seq 1
0.000.001 0.349.980 349.980"

ORIGIN
Query Match      100.0%; Score 753; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGCAGCCCTG 60
Db 276829 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCGCATCCAGAGCAGCCCTG 276888
QY 61 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 120
Db 276889 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 276948
QY 121 GCTGCAAGTTGGGCAATGACCCGATGATGCTTGGTGTGCTCTGCTTGGTTATTCATATAC 180
Db 276949 GCTGCAAGTTGGGCAATGACCCGATGATGCTTGGTGTGCTCTGCTTGGTTATTCATATAC 277008
QY 181 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTGATTTTGGCGGGCTCCACCGAA 240
Db 277009 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGCGCTGATTTTGGCGGGCTCCACCGAA 277068
QY 241 ATGCTGATCATGCGCCCTGTTGAGGCGGCAAGCCCTGAGGCGCATGCGCTCACCA 300
Db 277069 ATGCTGATCATGCGCCCTGTTGAGGCGGCAAGCCCTGAGGCGCATGCGCTCACCA 277128
QY 301 TTGCTGTGAATCTTCCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAA 360
Db 277129 TTGCTGTGAATCTTCCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAA 277188
QY 361 AACCCCATGCGCGTTTCTATTCGCTTTCGCGCTTATGAGAGCCATGACAGTCACT 420
Db 277189 AACCCCATGCGCGTTTCTATTCGCTTTCGCGCTTATGAGAGCCATGACAGTCACT 277248
QY 421 GCGGCAAGGCGCGAGGCTGTGCGGCTGCGGCTTATGCTCAATGCAATAGCGTTTAC 480
Db 277249 GCGGCAAGGCGCGAGGCTGTGCGGCTGCGGCTTATGCTCAATGCAATAGCGTTTAC 277308
QY 481 TCCACTGAGGTATTCGCGGCTGCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 540
Db 277309 TCCACTGAGGTATTCGCGGCTGCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 277368
QY 541 ATTAAGGCGCTCGAGTTTCGCGCTTTCGCTTTCGCTTTCGCTGACGCTGATTCCTGC 600
Db 277369 ATTAAGGCGCTCGAGTTTCGCGCTTTCGCTTTCGCTTTCGCTGACGCTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 660
Db 277429 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 277488
QY 661 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTTCTTGGGTCTGTGACC 720

```


gene	/db_xref="taxon:1718"	
	complement(385..711)	
	/gene="brnB"	
CDS	complement(385..711)	
	/gene="brnB"	
	/note="LIV-E transporter component 1; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnB"	
	/protein_id="AA04685.1"	
	/db_xref="GI:21311380"	
	/translation="MTDPSCLLVAVCAVTPALRAVPELILKPLRESQFVKMM MPAGTILALTRSTFRSNAIDKLTTFGLINAVLITVAHLGRRITLSVAGTIVFV GLVNLFF"	
gene	complement(708..1463)	
	/gene="DrmP"	
CDS	complement(708..1463)	
	/gene="DrmP"	
	/note="LIV-E transporter component 2; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnP"	
	/protein_id="AA04686.1"	
	/db_xref="GI:21311381"	
	/translation="MCKQDEHSSLEVSPSKALPPDKGYRYEIAQGLTSLAAGL GMYTIGLAGLVITQGYBMAAPLPSGLIRAGSTBMVIALVAGAPFLGIALTLL VNRHVFAPAFSPFLHVNKPIARFYSVALIDEAVVAPAAAPAGASARLISQIAFH SYWVFGGLTGAVALIIPFELKLEFALCSLVTLTLDSCRTKQIPELLAGHSFTI ALVVIPOGALFPALALIPLGLLIRYFPLGKAK"	
gene	1538..1993	
	/gene="Lrp"	
CDS	1538..1993	
	/gene="Lrp"	
	/note="Lrp"	
	/codon_start=1	
	/transl_table=11	
	/product="Lrp-like regulator"	
	/protein_id="AA04687.1"	
	/db_xref="GI:21311382"	
	/translation="MKLDSIDRAITAEISANARISNLAADKHLHTPSPCRRTORLE AGTILGYSADIHPAVMNKGFEVTVLISNDRSTVNFSSVAQHDVEVLEHLRFG SPDYFRIGVADLEAYEQFLSHIQTVGIAKISRFAMKVKVPAPQV"	
ORIGIN		
Query Match	100.0%; Score 753; DB 1; Length 2105;	
Best Local Similarity	100.0%; Pred. No. 1e-184;	
Matches 753; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCGCATCCAGGACAGCCCTG	60
DB	1463 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCGCATCCAGGACAGCCCTG	1404
QY	61 GAACCGATGATTAAGCTATGCGGCTGCGAATCGGACAGGTCTAAAACTCCCTT	120
DB	1403 GAACCGATGATTAAGCTATGCGGCTGCGAATCGGACAGGTCTAAAACTCCCTT	1344
QY	121 GTCGAGGTTTGGGCAATGATCCGATGGTATGGCTGTGCTCTTGGTTATTCATATC	180
DB	1343 GTCGAGGTTTGGGCAATGATCCGATGGTATGGCTGTGCTCTTGGTTATTCATATC	1284
QY	181 GGCTTACGATGTGGGACAGCCCACTGTTTTCGGGCTGATTTTGGCGGCTCCACCGAA	240
DB	1283 GGCTTACGATGTGGGACAGCCCACTGTTTTCGGGCTGATTTTGGCGGCTCCACCGAA	1224
QY	241 ATGCTGATCGCCCTGCTGTGTGGGACAGGCGCCCTGGGCGCATCGCGTCAACCA	300
DB	1223 ATGCTGATCGCCCTGCTGTGTGGGACAGGCGCCCTGGGCGCATCGCGTCAACCA	1164
QY	301 TTGCTGTGAATCTCCGCCAGATTCATTCAGGTTTCAATCCCGCTGCATGTGTCAAA	360
DB	1163 TTGCTGTGAATCTCCGCCAGATTCATTCAGGTTTCAATCCCGCTGCATGTGTCAAA	1104
QY	361 AACCCATTGCCCGTTTCTATTGCGTTTTCGGGCTTATCGACGAAGCTTACGAGTCACT	420
DB	1103 AACCCATTGCCCGTTTCTATTGCGTTTTCGGGCTTATCGACGAAGCTTACGAGTCACT	1044
QY	421 GCGGCGACGCGCGACAGCTGTGCGGCGGACGACTATCTCATGCAATGAAATAGGTTGAC	480
DB	1043 GCGGCGACGCGCGACAGCTGTGCGGCGGACGACTATCTCATGCAATGAAATAGGTTGAC	984
QY	481 TCCTACTGGTATTCGCGCGGTCTCACCGAGTGGCGCATCGAGATGATTCCTTTGAA	540
DB	983 TCCTACTGGTATTCGCGCGGTCTCACCGAGTGGCGCATCGAGATGATTCCTTTGAA	924
QY	541 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTCTTGTGACGCTGATTCGATTCGCG	600
DB	923 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTCTTGTGACGCTGATTCGATTCGCG	864
QY	601 CGAACGAAAAAGCAGATCCCTCTGCTGCTGCTGACAGTTTGACCTTCAACATTCGCTT	660
DB	863 CGAACGAAAAAGCAGATCCCTCTGCTGCTGCTGACAGTTTGACCTTCAACATTCGCTT	804
QY	661 GTGCTAATTCAGATCAGGCGCTTATTTGGGCGGCTGCTATCTTTGGGTCTGTTGAC	720
DB	803 GTGCTAATTCAGATCAGGCGCTTATTTGGGCGGCTGCTATCTTTGGGTCTGTTGAC	744
QY	721 ATCCGACTCTTCTTGGGAAAGGCTGCTAA	753
DB	743 ATCCGACTCTTCTTGGGAAAGGCTGCTAA	711
RESULT 13		
AP005274	340000 bp DNA linear BCT 08-AUG-2002	
LOCUS	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section	
DEFINITION	1/10.	
ACCESSION	AP005274 BA000036	
VERSION	AP005274.1 GI:21322764	
KEYWORDS		
SOURCE	Corynebacterium glutamicum ATCC 13032	
ORGANISM	Corynebacterium glutamicum ATCC 13032	
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.	
AUTHORS	Nakagawa, S.	
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 340000)	
AUTHORS	Nakagawa, S.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories, 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@xanaden.com, Tel.:81-44-829-3031, Fax:81-44-813-1651)	
COMMENT	This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.	
FEATURES	Location/Qualifiers	
source	1..340000	
gene	/organism="Corynebacterium glutamicum ATCC 13032"	
	/mol_type="genomic DNA"	
	/strain="ATCC 13032"	
	/db_xref="taxon:196627"	
CDS	/note="ATCC 13032"	
	1..1575	
	/gene="Cg10001"	
	1..1575	
	/note="PF00308: Bacterial dnaA protein	
	TRG00362: dnaA: chromosomal replication initiator protein	
	dnaA"	
	/codon_start=1	
	/transl_table=11	
	/product="ATPase involved in DNA replication initiation"	
	/protein_id="BAB97394.1"	

Db 581 TCTACTGGGATTCGCGGCTCTACCGAGTGGGATCGCAAGTTGATTCCTTTGAA 640
Qy 541 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTTTTGTCAAGCTGATTTGGATTCCTGC 600
Db 641 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTTTTGTCAAGCTGATTTGGATTCCTGC 700
Qy 601 CGAAGCAAAAGCAATCCCTCTCTGCTGCTCGAGGTTGAGCTTACCAATTCCTT 660
Db 701 CGAAGCAAAAGCAATCCCTCTCTGCTGCTCGAGGTTGAGCTTACCAATTCCTT 760
Qy 661 GTGGTAATTCAGTCAAGCCCTATTTGCGGCTGATCTTCTTGGATCTGTGACC 720
Db 761 GTGGTAATTCAGTCAAGCCCTATTTGCGGCTGATCTTCTTGGATCTGTGACC 820
Qy 721 ATCCGTAATCTTCTTGTGGAAAGCTGTAA 753
Db 821 ATCCGTAATCTTCTTGTGGAAAGCTGTAA 853
RESULT 11 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014994
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION BD014994
VERSION JP 2001169788-A/4.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kemerhnecht,N., Sahm,H., Eggerling,L. and Pfeifferle,W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL Patent: JP 2001169788-A 4 26-JUN-2001;
DEGUSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC13032
PN JP 2001169788-A/4
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KEMERHNECHT,HERMANN SAHM,LOTHAR EGGERLING,WALTER P.
PFEIFFERLE
PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/08//
PC C12N15/09,C12R1:15), (C12N1/21,C12R1:15), (C12P13/06,C12R1:15),
PC C12N15/00,
CC (C12N15/00,C12R1:15)
CC bnf
CC bnf
CC bnf
FH Key Location/Qualifiers
FT gene (101).. (853)
FT gene (853).. (1176).
FEATURES
source Location/Qualifiers
1..1271
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGCAAAAACGCAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCG 60
Db 101 GTGCAAAAACGCAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCG 160
Qy 61 GAACCAATGATTAAGGTTATCGCGCTACGAAATCGCGCAAGGTCTAAACCTCCCT 120
Db 161 GAACCAATGATTAAGGTTATCGCGCTACGAAATCGCGCAAGGTCTAAACCTCCCT 220
Qy 121 GCTGAGGTTGGGCACTACCGCATTCGATTCGCTTGTCTTGGTATTCAATAC 180

Db 221 GCTGAGGTTGGGCACTACCGCATTCGATTCGCTTGTCTTGGTATTCAATAC 280
Qy 181 GCTGAGGTTGGGCACTACCGCATTCGCTTGTCTTGGTATTCAATAC 240
Db 281 GCTGAGGTTGGGCACTACCGCATTCGCTTGTCTTGGTATTCAATAC 340
Qy 241 ATGCTGATGATCGCCCTCGTTGAGGCGAGGCGCCCTTGAGCGCATCGCTCAACA 300
Db 341 ATGCTGATGATCGCCCTCGTTGAGGCGAGGCGCCCTTGAGCGCATCGCTCAACA 400
Qy 301 TTGCTGATGATCGCCCTCGTTGAGGCGAGGCGCCCTTGAGCGCATCGCTCAACA 360
Db 401 TTGCTGATGATCGCCCTCGTTGAGGCGAGGCGCCCTTGAGCGCATCGCTCAACA 460
Qy 361 AACCCATGCGCGCTTCTATTGAGTTTTCGAGCTATGAGCAAGCTACAGCACT 420
Db 461 AACCCATGCGCGCTTCTATTGAGTTTTCGAGCTATGAGCAAGCTACAGCACT 520
Qy 421 GCGGCAAGGCGCGAGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 480
Db 521 GCGGCAAGGCGCGAGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 580
Qy 481 TCTACTGGGATTCGCGGCTGCTACCGAGTGGGATGCGAGATTGATTCCTTTGAA 540
Db 581 TCTACTGGGATTCGCGGCTGCTACCGAGTGGGATGCGAGATTGATTCCTTTGAA 640
Qy 541 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTTTGTGACGCTGATTCCTTGC 600
Db 641 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTTTGTGACGCTGATTCCTTGC 700
Qy 601 CGAAGCAAAAGCAATCCCTCTCTGCTGCTGCGAGTTGAGCTTACCAATTCCTT 660
Db 701 CGAAGCAAAAGCAATCCCTCTCTGCTGCTGCGAGTTGAGCTTACCAATTCCTT 760
Qy 661 GTGGTAATTCAGTCAAGCCCTATTTGCGGCTGATCTTCTTGGATCTGTGACC 720
Db 761 GTGGTAATTCAGTCAAGCCCTATTTGCGGCTGATCTTCTTGGATCTGTGACC 820
Qy 721 ATCCGTAATCTTCTTGTGGAAAGCTGTAA 753
Db 821 ATCCGTAATCTTCTTGTGGAAAGCTGTAA 853
RESULT 12
LOCUS AF454053/c 2105 bp DNA linear BCT 27-JUN-2002
DEFINITION Corynebacterium glutamicum Bnf (bnf), Bnf (bnf), and Lrp-like
regulator (Lrp) genes, complete cds.
ACCESSION AF454053
VERSION AF454053.1 GI:21311379
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 2105)
AUTHORS Kemerhnecht,N., Sahm,H., Yen,M.R., Patek,M., Saler Jr,M.H. Jr. and
Eggerling,L.
TITLE Export of L-isoleucine from Corynebacterium glutamicum: a
two-gene-encoded member of a new translocator family
JOURNAL J. Bacteriol. 184 (14), 3947-3956 (2002)
MEDLINE 22077265
PUBMED 12081967
REFERENCE 2 (bases 1 to 2105)
AUTHORS Kemerhnecht,N., Eggerling,L. and Sahm,H.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) IBT-1, Forschungszentrum, Leo-Brandt Str.,
Juelich 52425, Germany
FEATURES
source Location/Qualifiers
1..2105
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"

CDS
853..1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=11
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="REFSEQ:NC_003928"
/translation="MTDTSCLIVAVCAVTPALANPFLKLPRESQFGRKAM
MPAGILAILTASTFRSNAIDKTLTFGLIAVAITVAHLGGRRLTSLVAGTIVFV
GLVNLFF"

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGACAGATTCATTCAAGCCTGAGGAGTGTCCCATCCAGAGCCCTG 60
101 GTGCAAAAACGACAGATTCATTCAAGCCTGAGGAGTGTCCCATCCAGAGCCCTG 160
61 GAACCAAGTATTAAGTTATCGCGCTAGCAAAATCCGCGCAAGTCTTAAACCTCCCT 120
161 GAACCAAGTATTAAGTTATCGCGCTAGCAAAATCCGCGCAAGTCTTAAACCTCCCT 220
121 GCTGAGGTTGGGATGATCCGATTTGATTCGTTGGTCTCTTGTTATTCATATAC 180
221 GCTGAGGTTGGGATGATCCGATTTGATTCGTTGGTCTCTTGTTATTCATATAC 280
181 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACGAA 240
281 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACGAA 340
241 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
341 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
301 TTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
401 TTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 460
361 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 420
461 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 520
421 GCGGCAAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
521 GCGGCAAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
481 TCTTACTGGGATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
581 TCTTACTGGGATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 640
641 ATTAAAGGCTGAGATTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
701 ATTAAAGGCTGAGATTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
701 CGAAGCAAAACGACATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
661 GTGATTAATTCAGATTCAGGCTTATTTGCGGCGCTGATCTTCTTGGGCTGTGAGC 720
761 GTGATTAATTCAGATTCAGGCTTATTTGCGGCGCTGATCTTCTTGGGCTGTGAGC 820
721 ATCCGATCTTCT 753
821 ATCCGATCTTCT 853

RESULT 10
BD014991 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014991
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,

ACCESSION BD014991
VERSION BD014991.1 GI:2255798
KEYWORDS JP 2001169788-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kerner, R. H., Sahn, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
PATENT: JP 2001169788-A 1 26-JUN-2001;
DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/1
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KERNER, R. H. SAHN, LOTHAR EGGERING, WALTER PI
PFEIFFERLE
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08//
PC (C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
PC C12N15/00, C12R1:15)
CC brnE
CC brnE
CC brnE
FH Key (101) (853)
FT gene (853) (1176).
FEATURES
source
Location/Qualifiers
1..1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGACAGATTCATTCAAGCCTGAGGAGTGTCCCATCCAGAGCCCTG 60
101 GTGCAAAAACGACAGATTCATTCAAGCCTGAGGAGTGTCCCATCCAGAGCCCTG 160
61 GAACCAAGTATTAAGTTATCGCGCTAGCAAAATCCGCGCAAGTCTTAAACCTCCCT 120
161 GAACCAAGTATTAAGTTATCGCGCTAGCAAAATCCGCGCAAGTCTTAAACCTCCCT 220
121 GCTGAGGTTGGGATGATCCGATTTGATTCGTTGGTCTCTTGTTATTCATATAC 180
221 GCTGAGGTTGGGATGATCCGATTTGATTCGTTGGTCTCTTGTTATTCATATAC 280
181 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACGAA 240
281 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACGAA 340
241 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
341 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
301 TTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
401 TTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 460
361 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 420
461 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 520
421 GCGGCAAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
521 GCGGCAAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
481 TCTTACTGGGATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

Db 761 GTGTAATTCAGAGTCAGGCCCTATTGCGCGCTGCTGATCTTTGGGTCTTGACC 820

QY 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 853

RESULT 8

AX137709 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 1 from Patent EP1096010.

AX137709

ACCESSION AX137709.1 GI:14273886

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 Kernerhecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
Patent: EP 1096010-A 1 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

FEATURES

source

1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
101. 856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/translation="MOKTQEHSLSEVSPSKALEPDDKRYRREIAGLTSIAAGL
GMPPIGIAFGLVIOGYEWMAAPLPSGLIFASSTEMLVIALVGAAPLAIATLL
VNFHFVAFSPPLHVKNPILAFYSVPLIDEAVATAPAGSARLISMOIAFH
SYNFGILGVAIAELIPEIKIGLEFALCSFVTLTDSCKTKQIPSLIAGLSFTI
ALVVIPOALFAALILIFGLITIRYFLGRAX"

CDS

853. 1179
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/translation="MTTPDSCLIVAVCAVTFALRAVPLIKELAESOFVGMAM
WMPAGLIALITASTFRSNATIDKTLIFGLIAVITVVAHLGSRITLLSVAGTIVFY
GVNLP"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GTGCAAAAAACGACAGATTCATTCAAGCTCGAGGTGTCCGATCCAGGAGCCCTG 60

QY 101 GTGCAAAAAACGACAGATTCATTCAAGCTCGAGGTGTCCGATCCAGGAGCCCTG 160

Db 61 GAACCAAGATTAAGGTTATCGAGCTACGAAATCGCGAAGGTCTAAAAACCTCCCT 120

QY 161 GAACCAAGATTAAGGTTATCGAGCTACGAAATCGCGAAGGTCTAAAAACCTCCCT 220

Db 121 GCTGCAAGTTGGGCAATGTAACCGATTGATTGCTGCTCTTGGTATTCAATAC 180

QY 221 GCTGCAAGTTGGGCAATGTAACCGATTGATTGCTGCTCTTGGTATTCAATAC 280

Db 181 GGCTACGAAGTGGGAGCGGACGCTGTTCCGGGCTGATTTCCGGGCTCCACCGAA 240

QY 281 GGCTACGAAGTGGGAGCGGACGCTGTTCCGGGCTGATTTCCGGGCTCCACCGAA 340

QY 241 ATGCTGTCATATGCGCTGCTGTTGGGCGCAGGCCCCCGGCGCATGGGCTACACGA 300

Db 341 ATGCTGTCATGCGCTGCTGTTGGGCGCAGGCCCCCGGCGCATGGGCTACACGA 400

QY 301 TTGCTGTAACCTTCCGACAGTATTCATGCTTTGATTCCTCCGTCATGTCGCTAA 360

Db 401 TTGCTGTAACCTTCCGACAGTATTCATGCTTTGATTCCTCCGTCATGTCGCTAA 460

QY 361 AACCCATTCGCCGCTTCTTATTCGCTTTCCGCTTATTCGAGAGAGCTTACGATGCT 420

Db 461 AACCCATTCGCCGCTTCTTATTCGCTTTCCGCTTATTCGAGAGAGCTTACGATGCT 520

QY 421 GCGGCGAGGCGCGAGGCTGTGCGGCGTGGCGCATTCATCTCAATGCAATAGCGTTGAC 480

Db 521 GCGGCGAGGCGCGAGGCTGTGCGGCGTGGCGCATTCATCTCAATGCAATAGCGTTGAC 580

QY 481 TCCTACTGAGTATTCGCGCGCTTCAACCGAGTGGCGATTCGAGATTCCTTTTGA 540

Db 581 TCCTACTGAGTATTCGCGCGCTTCAACCGAGTGGCGATTCGAGATTCCTTTTGA 640

QY 541 ATTAAGGCGCTGAGTTCGCCCTTTCCTCTCTTGTGACGCTGATTCGATTCGCTGAC 600

Db 641 ATTAAGGCGCTGAGTTCGCCCTTTCCTCTCTTGTGACGCTGATTCGATTCGCTGAC 700

QY 601 CGAAGCAAAAAGCAGATTCCTTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCT 660

Db 701 CGAAGCAAAAAGCAGATTCCTTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCT 760

QY 661 GTGTAATTCAGAGTCAGGCCCTATTGCGCGCTGCTGATCTTTGGGTCTTGACC 720

Db 761 GTGTAATTCAGAGTCAGGCCCTATTGCGCGCTGCTGATCTTTGGGTCTTGACC 820

QY 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 853

RESULT 9

AX137714 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 6 from Patent EP1096010.

AX137714

ACCESSION AX137714.1 GI:14273893

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 Kernerhecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
Patent: EP 1096010-A 6 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

FEATURES

source

1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC13032"
101. 856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39927.1"
/db_xref="GI:14273894"
/translation="MOKTQEHSLSEVSPSKALEPDDKRYRREIAGLTSIAAGL
GMPPIGIAFGLVIOGYEWMAAPLPSGLIFASSTEMLVIALVGAAPLAIATLL
VNFHFVAFSPPLHVKNPILAFYSVPLIDEAVATAPAGSARLISMOIAFH
SYNFGILGVAIAELIPEIKIGLEFALCSFVTLTDSCKTKQIPSLIAGLSFTI
ALVVIPOALFAALILIFGLITIRYFLGRAX"


```

DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1271)
  Kennetknecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
  Nucleotide sequences coding for the export of branched chain amino
  acids, process for the isolation thereof and use thereof
  Patent: US 6613545-A 1 02-SEP-2003;
  Location/Qualifiers
    location=1..1271
    /organism="Unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match          100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGACCCCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGACCCCTG 160
QY 61 GAACCATGATTAAGATTACGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
DB 161 GAACCATGATTAAGATTACGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 220
QY 121 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 180
DB 221 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 280
QY 221 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 280
DB 181 GGCTACGAATGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGAA 240
DB 281 GGCTACGAATGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGAA 340
QY 241 ATGCTGATCATGCGCTCGTTGTGGGCGAGCGCCCTGGGGCCATCGCGCTCACCA 300
DB 341 ATGCTGATCATGCGCTCGTTGTGGGCGAGCGCCCTGGGGCCATCGCGCTCACCA 400
QY 301 TTGCTGATTAAGTCCGCAAGTATTCATGCGTTTCAATCCGCTGATGTGCTCAA 360
DB 401 TTGCTGATTAAGTCCGCAAGTATTCATGCGTTTCAATCCGCTGATGTGCTCAA 460
QY 361 AACCCCATTCGCCGTTTCTATTCGATTCGATTCGATTCGATTCGATTCGATTC 420
DB 461 AACCCCATTCGCCGTTTCTATTCGATTCGATTCGATTCGATTCGATTCGATTC 520
QY 421 GGGGCGAGGCGCGCAGGCTGCTGGGCGAGCTTATTCATATGCAATAGCGTTTAC 480
DB 521 GGGGCGAGGCGCGCAGGCTGCTGGGCGAGCTTATTCATATGCAATAGCGTTTAC 580
QY 481 TCTTACTGGGATTCGCGCGGCTCTCAACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
DB 581 TCTTACTGGGATTCGCGCGGCTCTCAACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 640
QY 541 ATTAAAGGCGCTGAGTTCGCCCTTGTGCTCTCTTTGTGCAAGTACCTTGGATTCCTGC 600
DB 641 ATTAAAGGCGCTGAGTTCGCCCTTGTGCTCTCTTTGTGCAAGTACCTTGGATTCCTGC 700
QY 601 CGAAGCAAAAAGCAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTACCATTCCTCT 660
DB 701 CGAAGCAAAAAGCAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTACCATTCCTCT 760
QY 661 GTGGTAAATTCAGTTCAGGCGCTTATTTGGCGGCTGCTGATCTTTGGGCTGTGAC 720
DB 761 GTGGTAAATTCAGTTCAGGCGCTTATTTGGCGGCTGCTGATCTTTGGGCTGTGAC 820
QY 721 ATCCGGTACTGCTGCTTGGGAAAGGCTGTAA 753
DB 821 ATCCGGTACTGCTTCTTGGGAAAGGCTGTAA 853

```

```

RESULT 7
AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1271)
  Kennetknecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
  Nucleotide sequences coding for the export of branched chain amino
  acids, process for the isolation thereof and use thereof
  Patent: US 6613545-A 6 02-SEP-2003;
  Location/Qualifiers
    location=1..1271
    /organism="Unknown"
    /mol_type="genomic DNA"
ORIGIN
Query Match          100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGACCCCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGTGTGCCATCCAGGACCCCTG 160
QY 61 GAACCATGATTAAGATTACGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 120
DB 161 GAACCATGATTAAGATTACGCGCTACGAATTCGCGCAAGTCTAAAACTCCCTT 220
QY 121 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 180
DB 221 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 280
QY 221 GCTGCAAGTTTGCGCATGTACCCGATGTGATTCGGTTGCTCTTGATTCAATAC 280
DB 181 GGCTACGAATGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGAA 240
DB 281 GGCTACGAATGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGAA 340
QY 241 ATGCTGATCATGCGCTCGTTGTGGGCGAGCGCCCTGGGGCCATCGCGCTCACCA 300
DB 341 ATGCTGATCATGCGCTCGTTGTGGGCGAGCGCCCTGGGGCCATCGCGCTCACCA 400
QY 301 TTGCTGATTAAGTCCGCAAGTATTCATGCGTTTCAATCCGCTGATGTGCTCAA 360
DB 401 TTGCTGATTAAGTCCGCAAGTATTCATGCGTTTCAATCCGCTGATGTGCTCAA 460
QY 361 AACCCCATTCGCCGTTTCTATTCGATTCGATTCGATTCGATTCGATTCGATTC 420
DB 461 AACCCCATTCGCCGTTTCTATTCGATTCGATTCGATTCGATTCGATTCGATTC 520
QY 421 GGGGCGAGGCGCGCAGGCTGCTGGGCGAGCTTATTCATATGCAATAGCGTTTAC 480
DB 521 GGGGCGAGGCGCGCAGGCTGCTGGGCGAGCTTATTCATATGCAATAGCGTTTAC 580
QY 481 TCTTACTGGGATTCGCGCGGCTCTCAACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
DB 581 TCTTACTGGGATTCGCGCGGCTCTCAACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 640
QY 541 ATTAAAGGCGCTGAGTTCGCCCTTGTGCTCTCTTTGTGCAAGTACCTTGGATTCCTGC 600
DB 641 ATTAAAGGCGCTGAGTTCGCCCTTGTGCTCTCTTTGTGCAAGTACCTTGGATTCCTGC 700
QY 601 CGAAGCAAAAAGCAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTACCATTCCTCT 660
DB 701 CGAAGCAAAAAGCAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTACCATTCCTCT 760
QY 661 GTGGTAAATTCAGTTCAGGCGCTTATTTGGCGGCTGCTGATCTTTGGGCTGTGAC 720

```



```

QY 181 GGCTACGAATGGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 240
DB 181 GGCTACGAATGGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 240
QY 241 ATGCTGATCATGCCCTCGTTGTGGGCGAGGCCCTGGGCGCCATCGCGTCAACCA 300
DB 241 ATGCTGATCATGCCCTCGTTGTGGGCGAGGCCCTGGGCGCCATCGCGTCAACCA 300
QY 301 TTGCTGTGAATCTTCCGCAAGTATCTATGCGTTTCACTTCCGCTGCAATGTGTCAA 360
DB 301 TTGCTGTGAATCTTCCGCAAGTATCTATGCGTTTCACTTCCGCTGCAATGTGTCAA 360
QY 361 AACCCATTTGCCGTTTCTATTGCGTTTGGCGTTATGACGAAAGCTTACAGTCACT 420
DB 361 AACCCATTTGCCGTTTCTATTGCGTTTGGCGTTATGACGAAAGCTTACAGTCACT 420
QY 421 GCGGCGAGCCCGCAGCGTGTGCGGCGAGCACTTATCTCAATGCAATAGCGTTTAC 480
DB 421 GCGGCGAGCCCGCAGCGTGTGCGGCGAGCACTTATCTCAATGCAATAGCGTTTAC 480
QY 481 TCCCTACTGGGATATTGGCGGCTGCAACGAGTGGGAGATGCGAGATGATTCCTTTGAA 540
DB 481 TCCCTACTGGGATATTGGCGGCTGCAACGAGTGGGAGATGCGAGATGATTCCTTTGAA 540
QY 541 ATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAGCGTGAATTGCTTGC 600
DB 541 ATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAGCGTGAATTGCTTGC 600
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
DB 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
QY 661 GTGGTAATTCAGAGTCAAGGCCCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACC 720
DB 661 GTGGTAATTCAGAGTCAAGGCCCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACC 720
QY 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
DB 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 5
BD162490 753 bp DNA linear PAT 17-JAN-2003
LOCUS BD162490 Novel polynucleotide.
DEFINITION BD162490
ACCESSION BD162490.1 GI:27868248
VERSION JP 2002191370-A/289.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 753)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayaashi,M., Ochiai,K.,
TITLE Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
JOURNAL Novel polynucleotide
Patent: JP 2002191370-A 289 09-JUL-2002;
COMMENT KYOWA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/289
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
PC 04,C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
G01N33/56,
PC G01N33/569,G01N33/68,G01N37/00/C12P21/08,C12N1/21,C12R1/15,
PC C12N1/21,C12R1/13,C12N1/21,C12R1/01,C12P13/08,C12R1/15,

```

```

FEATURES
    source
        1..753
        /organism="unclassified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGGTGTCCCATTCAGAGGCGCTG 60
DB 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGGTGTCCCATTCAGAGGCGCTG 60
QY 61 GAACCAATATTAAGTTATCCGCGCTACGAATCGCGAGGCTTAAAACTCCCTT 120
DB 61 GAACCAATATTAAGTTATCCGCGCTACGAATCGCGAGGCTTAAAACTCCCTT 120
QY 121 GCTGAGGTTTGGCATGACCCGATTGATTCGTTTGTCTTGTGTTATTCAATAC 180
DB 121 GCTGAGGTTTGGCATGACCCGATTGATTCGTTTGTCTTGTGTTATTCAATAC 180
QY 181 GGTACGAATGTGGGAGGCCCACTGTTTCCGCGCTGATTTGGCGGCTCCACCGAA 240
DB 181 GGTACGAATGTGGGAGGCCCACTGTTTCCGCGCTGATTTGGCGGCTCCACCGAA 240
QY 241 ATGCTGATTCAGGCTGCTGTTGTGGGCGAGCGCCCTGGGCGCCATCGGCTCACACA 300
DB 241 ATGCTGATTCAGGCTGCTGTTGTGGGCGAGCGCCCTGGGCGCCATCGGCTCACACA 300
QY 301 TTGCTGTGAATCTTCCGCAAGTATCTATGCGTTTCACTTCCGCTGCAATGTGTCAA 360
DB 301 TTGCTGTGAATCTTCCGCAAGTATCTATGCGTTTCACTTCCGCTGCAATGTGTCAA 360
QY 361 AACCCATTTGCCGTTTCTATTGCGTTTGGCGTTATGACGAAAGCTTACAGTCACT 420
DB 361 AACCCATTTGCCGTTTCTATTGCGTTTGGCGTTATGACGAAAGCTTACAGTCACT 420
QY 421 GCGGCGAGCCCGCAGCGTGTGCGGCGAGCACTTATCTCAATGCAATAGCGTTTAC 480
DB 421 GCGGCGAGCCCGCAGCGTGTGCGGCGAGCACTTATCTCAATGCAATAGCGTTTAC 480
QY 481 TCCCTACTGGGATATTGGCGGCTGCAACGAGTGGGAGATGCGAGATGATTCCTTTGAA 540
DB 481 TCCCTACTGGGATATTGGCGGCTGCAACGAGTGGGAGATGCGAGATGATTCCTTTGAA 540
QY 541 ATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAGCGTGAATTGCTTGC 600
DB 541 ATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTCAGCGTGAATTGCTTGC 600
QY 601 GAAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTCAACATTCGCTT 660
DB 601 GAAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTCAACATTCGCTT 660
QY 661 GTGGTAATTCAGAGTCAAGGCCCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACC 720
DB 661 GTGGTAATTCAGAGTCAAGGCCCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACC 720
QY 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
DB 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 6
AR391953 1271 bp DNA linear PAT 18-DEC-2003
LOCUS AR391953

```

LOCUS AX137710 753 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 2 from Patent EP1096010.
ACCESSION AX137710
VERSION AX137710.1 GI:14273889
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinomycetales;
Corynebacteriaceae; Corynebacteriales; Corynebacterium.
REFERENCE
AUTHORS Kernerkecht,N., Eggeling,L., Sahm,H. and Pfeifferle,W.
TITLE Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 2 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
1..>753
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39925.1"
/db_xref="GI:14273890"
/db_xref="EMBL:CA39925"
/translation="MOKQEHHSLEVSPSRAALEPDDKGRYREINAGLTSLAAGL
VNFHFVAFSPFLHVKPILKRFYSVALIDEAYATAAPAGMSARLISKOIARH
SLYVFGGLGVAIAELIPEIKGLFELFACSLFVTLTLDSCRTKQIPSLHAGSFTI
ALVIVPGALFPAALIFLGLTIRYFLEKAK"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGACAGATTATTCATTCAGCTGAGGTGTCGCATCCAGGACGCTG 60
DB 1 GTGCAAAAAACGACAGATTATTCATTCAGCTGAGGTGTCGCATCCAGGACGCTG 60
QY 61 GAACCGATGATTAAGATTATGCGGCTCGAATTCGCGGCAAGTCTAAACCTCCCTT 120
DB 61 GAACCGATGATTAAGATTATGCGGCTCGAATTCGCGGCAAGTCTAAACCTCCCTT 120
QY 121 GGTGCAAGTTGGGCAATGATACCGGATGATGATGATGATGATGATGATGATGATG 180
DB 121 GGTGCAAGTTGGGCAATGATACCGGATGATGATGATGATGATGATGATGATGATG 180
QY 181 GGTGCAAGTTGGGCAAGTCTGTTTCCGCTGATTTTGGCGGCTCCACCGAA 240
DB 181 GGTGCAAGTTGGGCAAGTCTGTTTCCGCTGATTTTGGCGGCTCCACCGAA 240
QY 241 ATGCTGATGATGCGCTGCTGTTGGGCGGACGCGCTGAGGCGCATCGGCTCACCA 300
DB 241 ATGCTGATGATGCGCTGCTGTTGGGCGGACGCGCTGAGGCGCATCGGCTCACCA 300
QY 301 TTGCTGATGATGCGCTGCTGTTGGGCGGACGCGCTGAGGCGCATCGGCTCACCA 360
DB 301 TTGCTGATGATGCGCTGCTGTTGGGCGGACGCGCTGAGGCGCATCGGCTCACCA 360
QY 361 AACCCATGATGCGCTGCTGTTGGGCGGCTGATGATGATGATGATGATGATGATGAT 420
DB 361 AACCCATGATGCGCTGCTGTTGGGCGGCTGATGATGATGATGATGATGATGATGAT 420
QY 421 GCGGCGAGGCGCGGCGGCTGCTGTTGGGCGGCTGATGATGATGATGATGATGATGAT 480
DB 421 GCGGCGAGGCGCGGCGGCTGCTGTTGGGCGGCTGATGATGATGATGATGATGATGAT 480
QY 481 TCTACTGGGATTTGGCGGCTGCTGCGGAGTGGGAGTGGAGTGGATGATGATGATGATG 540
DB 481 TCTACTGGGATTTGGCGGCTGCTGCGGAGTGGGAGTGGAGTGGATGATGATGATGATG 540

DB 481 TCTACTGGGATTTGGCGGCTGCTGCGGAGTGGGAGTGGAGTGGATGATGATGATGATG 540
QY 541 ATTAAGGCGCTGAGATTGCGCCCTTGGCTCTCTTTGTCAGCGTGAATTGCTGCG 600
DB 541 ATTAAGGCGCTGAGATTGCGCCCTTGGCTCTCTTTGTCAGCGTGAATTGCTGCG 600
QY 601 CGAACGAAAAAGAGATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 CGAACGAAAAAGAGATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 GTGTAATTCAGATGAGGCGCTTATTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 GTGTAATTCAGATGAGGCGCTTATTTGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTAA 753
DB 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTAA 753
RESULT 4
BD014992 753 bp DNA linear PAT 27-AUG-2002
LOCUS BD014992
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION BD014992
VERSION BD014992.1 GI:22555799
KEYWORDS JP 2001169788-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
AUTHORS Kernerkecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL Patent: JP 2001169788-A 2 26-JUN-2001;
Degussa HUELIS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT
OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/2
PD 26-JUN-2001 JP 2000324315
PF 24-OCT-2000 JP 19951708.8
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KERNERKECHT, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
PFEIFFERLE
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC C12N15/09, C12R1.15), (C12N1/21, C12R1.15), (C12P13/06, C12R1.15),
PC C12N15/00, C12R1.15)
CC brnF
CC key
FH key
FT CDS location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGACAGATTATTCATTCAGCTGAGGTGTCGCATCCAGGACGCTG 60
DB 1 GTGCAAAAAACGACAGATTATTCATTCAGCTGAGGTGTCGCATCCAGGACGCTG 60
QY 61 GAACCGATGATTAAGATTATGCGGCTCGAATTCGCGGCAAGTCTAAACCTCCCTT 120
DB 61 GAACCGATGATTAAGATTATGCGGCTCGAATTCGCGGCAAGTCTAAACCTCCCTT 120
QY 121 GGTGCAAGTTGGGCAATGATACCGGATGATGATGATGATGATGATGATGATGATG 180
DB 121 GGTGCAAGTTGGGCAATGATACCGGATGATGATGATGATGATGATGATGATGATG 180

FEATURES
source
Location/Qualifiers
1..753
/organism="unknown"
/mol_type="genomic DNA"

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
61 GAACCAATGATTAAGATTATGGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 120
61 GAACCAATGATTAAGATTATGGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 120
121 GCTGACAGTTGGGACATGATACCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
121 GCTGACAGTTGGGACATGATACCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
181 GGTACGATGATGGGAGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 240
181 GGTACGATGATGGGAGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 240
241 ATGCTGTGATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 ATGCTGTGATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 TTGCTGTGATGCTGCGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
301 TTGCTGTGATGCTGCGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
361 AACCCATTGCGCGCTTCTATTCGCTTTTTCGCGCTTATCGAGCAAGCTTACGATCA 420
361 AACCCATTGCGCGCTTCTATTCGCTTTTTCGCGCTTATCGAGCAAGCTTACGATCA 420
421 GCGGCAAGCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
421 GCGGCAAGCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
481 TCTCTAGGATTCGCGGCTTCAACCGGAGTGGCAATGCGAGTTGATTCCTTTGAA 540
481 TCTCTAGGATTCGCGGCTTCAACCGGAGTGGCAATGCGAGTTGATTCCTTTGAA 540
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTTTGTCAGCTGATCTTGTGATTCCTG 600
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTTTGTCAGCTGATCTTGTGATTCCTG 600
601 CGAAGCAAAAAACGATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
601 CGAAGCAAAAAACGATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
661 GTGTAATTCAGGCTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
661 GTGTAATTCAGGCTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
721 ATCCGATCTTCT 753
721 ATCCGATCTTCT 753

RESULT 2
AX120373 753 bp DNA linear PAT 11-MAY-2001
LOCUS AX120373
DEFINITION Sequence 289 from Patent EP1108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE
AUTHORS
TITLE
JOURNAL
Kyowa Hakko Kogyo Co., Ltd. (JP)
Patent: EP 1108790-A 289 20-JUN-2001;
Novel polynucleotides
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,

FEATURES

source
Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
61 GAACCAATGATTAAGATTATGGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 120
61 GAACCAATGATTAAGATTATGGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 120
121 GCTGACAGTTGGGACATGATACCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
121 GCTGACAGTTGGGACATGATACCGGATTTGCTGCTGCTGCTGCTGCTGCTGCTG 180
181 GGTACGATGATGGGAGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 240
181 GGTACGATGATGGGAGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 240
241 ATGCTGTGATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 ATGCTGTGATGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 TTGCTGTGATGCTGCGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
301 TTGCTGTGATGCTGCGCAGATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
361 AACCCATTGCGCGCTTCTATTCGCTTTTTCGCGCTTATCGAGCAAGCTTACGATCA 420
361 AACCCATTGCGCGCTTCTATTCGCTTTTTCGCGCTTATCGAGCAAGCTTACGATCA 420
421 GCGGCAAGCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
421 GCGGCAAGCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
481 TCTCTAGGATTCGCGGCTTCAACCGGAGTGGCAATGCGAGTTGATTCCTTTGAA 540
481 TCTCTAGGATTCGCGGCTTCAACCGGAGTGGCAATGCGAGTTGATTCCTTTGAA 540
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTTTGTCAGCTGATCTTGTGATTCCTG 600
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTTTGTCAGCTGATCTTGTGATTCCTG 600
601 CGAAGCAAAAAACGATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
601 CGAAGCAAAAAACGATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
661 GTGTAATTCAGGCTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
661 GTGTAATTCAGGCTCAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
721 ATCCGATCTTCT 753
721 ATCCGATCTTCT 753

RESULT 3
AX137710

Tue Apr 20 06:47:24 2004

us-10-608-504-6_copy_101_853.rge

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 / Search time 3191.51 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753
Sequence: 1 ggcacaaacgcagagat.....cttcggagaaagcgcctctaaa 753

Scoring table: IDENTITY_NTC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

GenBdb1:
1: gb_ba:*
2: gb_ba:*
3: gb_ba:*
4: gb_ba:*
5: gb_ba:*
6: gb_ba:*
7: gb_ba:*
8: gb_ba:*
9: gb_ba:*
10: gb_ba:*
11: gb_ba:*
12: gb_ba:*
13: gb_ba:*
14: gb_ba:*
15: gb_ba:*
16: gb_ba:*
17: gb_ba:*
18: gb_ba:*
19: gb_ba:*
20: gb_ba:*
21: gb_ba:*
22: gb_ba:*
23: gb_ba:*
24: gb_ba:*
25: gb_ba:*
26: gb_ba:*
27: gb_ba:*
28: gb_ba:*
29: gb_ba:*
30: gb_ba:*
31: gb_ba:*
32: gb_ba:*
33: gb_ba:*
34: gb_ba:*
35: gb_ba:*
36: gb_ba:*
37: gb_ba:*
38: gb_ba:*
39: gb_ba:*
40: gb_ba:*
41: gb_ba:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	AR391954	AR391954 Sequence
2	753	100.0	753	AX120373	AX120373 Sequence
3	753	100.0	753	AX137710	AX137710 Sequence
4	753	100.0	753	BD014992	BD014992 Nucleotide
5	753	100.0	753	BD162490	BD162490 Novel pol
6	753	100.0	1271	AR391953	AR391953 Sequence
7	753	100.0	1271	AR391956	AR391956 Sequence
8	753	100.0	1271	AX137709	AX137709 Sequence
9	753	100.0	1271	AX137714	AX137714 Sequence
10	753	100.0	1271	BD014991	BD014991 Nucleotide
11	753	100.0	1271	BD014994	BD014994 Nucleotide
12	753	100.0	1271	AF454053	AF454053 Coryneb
13	753	100.0	2105	AP005274	AP005274 Coryneb
14	753	100.0	349980	AX120085	AX120085 Sequence
15	273.6	36.3	397	AX137083	AX137083 Sequence
16	273.6	36.3	778	AX137085	AX137085 Sequence
17	149.6	19.9	87340	AP005224	AP005224 Coryneb
18	132.6	17.6	53793	AX248361	AX248361 Coryneb
19	127.4	16.9	711	AX123492	AX123492 Sequence
20	127.4	16.9	711	BD165609	BD165609 Novel pol
21	127.4	16.9	309400	AX127153	AX127153 Sequence
22	127.4	16.9	325651	AP005283	AP005283 Coryneb
23	85	11.3	715	AX137077	AX137077 Sequence
24	85	11.3	715	BD013995	BD013995 Novel nuc
25	59.8	7.9	349926	AX571660	AX571660 Wolinella
26	58.6	7.8	756	AR366287	AR366287 Sequence
27	57.2	7.6	9888	AE011050	AE011050 Methanosa
28	54.4	7.2	856	HSB325448	HSB325448 Homo sapi
29	51.8	6.9	11025	AE013585	AE013585 Methanosa
30	51	6.8	308015	AE016783	AE016783 Pseudomon
31	48	6.4	301838	AE017209	AE017209 Geobacter
32	47	6.2	2000	AX655393	AX655393 Sequence
33	46.6	6.2	3619	ECR410307	ECR410307 Erythrina c
34	46.6	6.2	10733	AE013695	AE013695 Versinia
35	46.6	6.2	220050	AX114156	AX114156 Versinia
36	46.2	6.1	7218	166494	166494 Sequence 14
37	46	6.1	215050	AL646084	AL646084 Ralstonia
38	45.8	6.1	20342	AE008888	AE008888 Salmonella
39	45.4	6.0	348624	AX640441	AX640441 Bordetella
40	45.4	6.0	348666	AX640426	AX640426 Bordetella
41	44.6	5.9	329709	AP002997	AP002997 Mesorhiz
42	44.2	5.9	208524	AE016825	AE016825 Chromobac
43	44.2	5.9	264050	AF627279	AF627279 Salmonella
44	44.2	5.9	300592	AE016846	AE016846 Salmonella
45	43.4	5.8	346294	AP002999	AP002999 Mesorhiz

ALIGNMENTS

RESULT 1	AR391954	AR391954	753 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR391954	Sequence 2 from patent US 6613545.				
DEFINITION	AR391954	Sequence 2 from patent US 6613545.				
ACCESSION	AR391954	Sequence 2 from patent US 6613545.				
VERSION	AR391954.1	GI:40115725				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 753)					
AUTHORS	Kennethnecht,N., Sahm,H., Eggeling,I., and Pfeifferle,W.					
TITLE	Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof					
JOURNAL	Patent: US 6613545-A 2 02-SEP-2003;					

Oy	119	ATTCATTCAAGCCCTGGAGAAGTGTGCGGCATTCGAAGCACCCTCGAACAACATGATTAAGGT	178
Db	825	MDDMMNNNNKNDNRKKKNKKKAKKKKKKKDKKDMDKDMKKKKAMKMAKKK	884
Oy	179	TATCGGCGCTACGAATTCGCGCAGTCTAAAACTCCCTTGTCGAGGTTGGGCATG	238
Db	885	KKKKKDAKKKKMFKKKKKAAAKKKKKKKKMNMNHNKKMKKKKKDMMNKMN	944
Oy	239	TACCGAATGGTATTTGCGTTTGATCTCTGGTATTCATAACGCGCTACGAATGGTGCGCA	298
Db	945	KKKKKKKKTKKKTTTTTTTTTTTTHTMTMMNNHNTTTTTTMTTTTTHHKKKKKKKKK	1004
Oy	299	GCCCCACTGTTTTCCGSGCTGATTTTGGCGGGCTCACCGAAATGCTGTCATCGGCCCTC	358
Db	1005	KNDNMNNMMKKKKKKNNKKKKKKKKKKKKNNKKKKKKNNNNNNMMKKKKMMNNKK	1064
Oy	359	GTTCGGCGCAGCGCCCCTGGGCGCATTCGCGCTACACATTTGCTGGTAACCTTCGCG	418
Db	1065	KKKKKKKCCCBKBCCCCCKKYKCBBKKYCYKCKKKKKKKKKYCCBCCCKKKCBGYCC	1124
Oy	419	CACGATTTCTATGCGTTTCAAT	441
Db	1125	CCCKKKRKBCKKKKKKCCCKKK	1147

RESULT 15	CB657749	815 bp	mRNA	linear	EST 09-APR-2003
LOCUS	CB657749				
DEFINITION	OSUNECC13F02.f OSUNECC Oryza sativa (Japanica cultivar-group) cDNA clone OSUNECC13F02 5', mRNA sequence.				

ACCESSION	CB657749
VERSION	CB657749.1
KEYWORDS	GI:29661474
SOURCE	EST.
ORGANISM	<i>Oryza sativa</i> [japonica cultivar-group]
	<i>Oryza sativa</i> [japonica cultivar-group]

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 835)	Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E., Kudrna, D., Dean, R., Soderlund, C., Wigg, R., and Wang, G.	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe oryzae	Unpublished (2003)

COMMENT
Contact: Rod Wang
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA

```

Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat ggc cat g
Plate: 13 row: F column: 02
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
FEATURES

```

```
source
1..835
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
```

```

/c1one="OSJNEc1JF03"
/r1ssue_type="leaf"
/dev_stage="3 week"
/lab_host="DH103"
/c1one_lib="OSJNEc"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

```

```

Query Match          3.0%; Score 38; DB 14; Length 835;
Best Local Similarity 51.1%; Pred. No. 28;
Matches      89; Conservative      0; Mismatches  85; Indels      0; Gaps      0;

Qy      352 CGCCCTGTTGTGGCGCGAGCGCCCTTGGGCGCATGGCGGTACACGATTTGCTGGTGA 411
Db      79  CGACGCGCTCTCGGTGCGCGGCGCTCGGGTCTTGGCGGCGCGGAGATGGCTTC 138

Qy      412 CTTCGCGACGATATCTATGCGTTTCACTCCGCGTGCATGTGTCAAAAACCCCATTGC 471
Db      139 ATGCCGCGGGGTTCGACCGGATGGCGCACCCGACCTGCGGCGCTGGAACGCGTGGCT 198

Qy      472 CGGTTCTATTGCGTTTCGCGGCTTATGAGAGAGCCACGCAATCACTGGGG 525
Db      199 CTCGCGCTACGCGCGCTCTTCGTGCGAGGAGCGTGGCGCCGCACTTCTCGGC 252

Search completed: April 17, 2004, 20:01:48
Job time : 416.78 secs

```

D_b 199 CTCGCCTACGCGGCTCTTCGTGCCAGGAGCAGTGGCCCTCGCCACTTCGGCGC 252

Search completed: April 17, 2004, 20:01:48
Job time : 4116.78 secs

http://image.llnl.gov
Plate: LLM9410 row: K column: 05
High quality sequence stop: 385.
Location/Qualifiers

FEATURES

source

1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4149244"
/issue_type="Globalatoma with EGFR amplification"
/lab_host="MD10B (T1 phase-resistant)"
/clone_lib="NCI CGAP Brn4"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 3.1%; Score 39; DB 10; Length 859;
Best Local Similarity 51.4%; Pred. No. 15;
Matches 90; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

760 TGTGTAATTCAGGTCAGCCCTATTGCGGCGCTGATCTTGGTCTGTAC 819
792 TGTGTTCTTCGGGTGTCCTTCTGTCGCCCGTCTCTGTTTCTGCG 733
820 CATCCGTAATCTTCTTGGGAAAGCGTCTTAATGACACTGATTTCTCTC 879
732 TCTGATATATTCCTTCTGCTGCTGCGTCTTCTTCTTCTTCTTCTG 673
QY 880 CTCTTGTGCGAGTATGCGAGCATTTTGTGCGCGCGGCGGTTCCGTTCT 934
DB 672 GTTGTGCTGCTGATGCTGCTGCTTGTGTTATCTCTGCTGCTGCTGCT 618

RESULT 13
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TERT end of BAC #
DEFINITION BACR19D16 of RCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION CNS0091P
VERSION AL053013 GI:4934461
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Karuoto Oosagawa and Aaron Mammeter in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; on bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bdpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"

FEATURES

source

1..925
/organism="Drosophila melanogaster"

/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TERT"

ORIGIN

Query Match 3.1%; Score 39; DB 29; Length 925;
Best Local Similarity 14.2%; Pred. No. 15;
Matches 37; Conservative 117; Mismatches 106; Indels 0; Gaps 0;

295 GGACAGCCCACTTTTCCGCTGATTTTCGGGCGCTACACCAATGCTGATCCG 354
638 SSSSSSTSSSTSSSTSSSSSSSSSSSSSSSTTSSTASGSGMSAGGSSSTSTSS 697
355 CTTGTTGGGGCGAGCGCCCTGGCGCCATCGGCTCACCACTTCTGTGACTT 414
698 SSSSSSTSSSTSSSVSSGSKSTBSGSSSSSSSTSSBBSCTSTSSSSSVSSST 757
QY 415 CCGCCAGCTATCTTACGCTTTCATTCGCCGTCAGTGTCAAAACCCCATTCGCG 474
DB 758 CCGCTCCCSYSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSST 817
QY 475 TTCTATTCGATTTTCGCGCTTATACAGAGCCTACGACGCTACGCGCCAGCGCCG 534
DB 818 YBMGCTSTGCGSSSSSSGKGVTKCGCGCGSSSTNGMBGTSSACSSSSSSSVSS 877
QY 535 AGCGTGTGCGGCTGCGGAC 554
DB 878 SSKSSASSSVSSGSSGS 897

RESULT 14

LOCUS BX361080 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION BX361080
VERSION BX361080.1 GI:30374504
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J., and Polayres, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr; Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue, Genoscope, sequence ID : CSOD1079D08P1.
Location/Qualifiers
1..1201

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1079P16"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dr) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.0%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 10.5%; Pred. No. 22;
Matches 34; Conservative 147; Mismatches 142; Indels 0; Gaps 0;

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
AL054013	AL054013.1	GI:4935590	SS.	Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
ACCESSION	AL054013	AL054013.1	GI:4935590	Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
VERSION	AL054013	AL054013.1	GI:4935590	Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
KEYWORDS	SS.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
SOURCE	Drosophila melanogaster (fruit fly)			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
REFERENCE	1 (bases 1 to 787)			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
AUTHORS	Genoscope.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
TITLE	Direct Submission			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RP11-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
FEATURES	Location/Qualifiers			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
SOURCE	1..787			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/organism="Drosophila melanogaster"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/mol_type="genomic DNA"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/db_xref="taxon:7227"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/clone="BACR20K04"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/clone_1lb="RP11-98"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	/note="end : T7"			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
ORIGIN				Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Query Match	3.1%; Score 39.4; DB 29; Length 787;			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Best Local Similarity	45.5%; Pred. No. 11;			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Matches	76; Conservative 15; Mismatches 76; Indels 0; Gaps 0;			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Y	652 CGAGTTGGCCTTGGCTCTCTCTTGTGACGCTACTTGGATTCCGCGACGAAAA 711			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Db	334 CTACTTGGCTCTCTCTCTCTTGTGTTGTAAMGMMGMMCTTGAATCCCTKXAGAPRAC 275			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Y	712 GCAGATCCCTCTCTGTGCTCGAGGTTTGAGCTTCACATTGCTCTGTGTAATCC 771			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Db	274 AYTGTGTTKTTTTCYACMGRCCTRAAGGTAAATTTCTTAATTTTGTGTTAAATTT 215			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Y	772 AGGTGAGCCCTATTGTTGGCGCGCTGCTGATCTTCTTGGGTCTGTGA 818			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
Db	214 TAATTGGCCATATGTTATATCGAATTTTATTTATGTCGCGTGA 168			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
RESULT 10				Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
LOCUS	BX425603 885 bp mRNA linear EST 15-MAY-2001			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
DEFINITION	BX425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
ACCESSION	CL08B022ZJ07 3-PRIME, mRNA sequence.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
VERSION	BX425603			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
KEYWORDS	BX425603.1 GI:30770486			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
SOURCE	EST.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
ORGANISM	Homo sapiens (human)			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
AUTHORS	Li, M.B., Gruber, C., Jesssee, J. and Polayes, D.			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
TITLE	Full-length cDNA libraries and normalization			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			
JOURNAL	Unpublished (2001)			Drosophila melanogaster (fruit fly)	fly), genomic survey sequence.			

```
organism="Gallus gallus"  
/mol_type="mRNA"  
/strain="CB"  
/db_xref="taxon:9031"  
/clone="241y1"  
tissue_type="Bursa of Fabricius"  
cell_type="Bursal lymphocyte"  
dev_stages="2-3 weeks old"
```

Db	515	GTTCGCGGCGGCGCCACCCGGAATCCAGCTCGCTTCGCGAGCGCCTCTAAAGCTTCAGGTC	594
Oy	651	TCGAGTTCGCCCTTTGCTCTCTCTTTGTCAGCG	683
Db	595	ACGAGAGGAGCGCTTTGCTCTCTCTCCCTGTCGCG	627

[illegible]

ORIGIN

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized.

Query Match	3.28; Score 40.8; DB 13; Length 1201;
Best Local Similarity	3.04; Pred. No. 5.2;
Matches 17; Conservative 191; Mismatches 362; Indels 0; Gaps 0;	
397 CACATTGCTGGTGAACATCCGCCACGATCTCTATGGCTTTCAATCCCGCTGCATGGT	456
406 CKMMNNNNKKKKCKMNNKKNNKKNNKKNNKKNNKKNNKKNNKKNNKKNNKKNNKKNN	465
457 CAAAAACCCATGCGCCGTTCTATTGGCTTTTCGCGCTTATGACAGAACCTACGAGT	516
466 MNN	525
517 CACTGCGGCGGAGCGCCGAGCGTGGTGGGGTGGCGACTTATCTCAATGCAATAGCGTT	576
526 MNNNNCKMMNN	585
577 TCACTCTACTGGGTATTCGCGCGTCTCACCGAGTGGCGAGATGGATGATTCCTT	636
586 NNN	645
637 TGAATTAAGGCGCTCGAGTTCGCGCTTGGCTCTCTTTGGTACGCTGACTTGATTC	696
646 KNN	705
697 CTGCGGAACGAAAAAGCAGATCCCTCTCTGCTGCTCGCAGGTTGACCTTCAACATTC	756
706 NKKNNKTKTKKKKKNAKKNNNTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	765
757 TCTTGATGTAATTCAGGTCAGGCGCTTATTCGAGCGCTGTGATCTTCTTGCTGTT	816
766 KNN	825
817 GACCATCCGGAATCTTCTTGGGAAAGGCTGCTAAATGACAATGATTTCTCCTGATTT	876
826 KCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKMCKM	885
877 TCTCTGTGTGCGAGTATGCGAGTCTTATTTGGCTCGCGGCGGTTCCGTTCTTA	936
886 MNN	945
937 ATCCTTAAGCCCTTACGATGATCAATTT 966	
946 MNNKCVNNKKNNKNNCKKKCVNNKKNNCKKKKK 975	

RESULT 5	692 bp	mrna	linear	EST 28-JUL-2003
BX540413/c				
LOCUS	BX540413			
DEFINITION	BX540413 AGENAE Gallus gallus multi-tissue library (gcab)			Gallus
ACCESSION	BX540413			
VERSION	BX540413.1			
KEYWORDS	EST			
SOURCE	Gallus gallus (chicken)			
ORGANISM	Gallus gallus			

REFERENCE	1 (bases 1 to 692)
AUTHORS	Herault, P., Le Mounh-Meuninger, V., Desert, C., Recout, E., Plum, F., Klupp, C. and Douaïre, M.
TITLE	Construction and primary characterization of chicken normalized multi-tissue cDNA libraries
JOURNAL	Unpublished (2003)
COMMENT	Contact: Douaïre M INRA, UMR INRA-ENSAR Genetique Animale 65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE Tel: +33 (0) 2.23.48.54.63

Email: Raf.Podowski@agr.ki.se
 Class: NotI site
 Location/Qualifiers
 source

1.473
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone_lib="Human NotI clones"

ORIGIN

Query Match 4.3%; Score 54.6; DB 28; Length 473;
 Best Local Similarity 56.2%; Pred. No. 0.0036;
 Matches 99; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 293 TGGGACAGCCCACTGTTTTCGGCCGATTTTCGGGGGCTCCACGAATGCTGTCATC 352
 DB 283 TGGGAGGGGACAGGGGCTTCGGCCATGCTTCGGGGGCGCGCCAGTTGTTGCCATC 342
 QY 353 GGCCTCGTTGTGGGCGACAGCCGCTGGGGCCATCGGCTCACCACTTGTGTGAAC 412
 DB 343 GGCATGCTTAAAGGGCGACCACTGCTGTGATCTGTACACCACTGCTGTGACT 402
 QY 413 TTCCGCGACGATTCATTCGCTTTTCATCCCGCTGACATGTCGAAAACCCCAT 468
 DB 403 TCGCAGACCTGCTTACGCGCTTGTCATGCGCCCGCTGCTGTACAGCACCCTT 458

RESULT 2
 BX356664 1201 bp mRNA linear EST 05-MAY-2003
 LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1015YB03 3-PRIME, mRNA sequence.
 ACCESSION BX356664
 VERSION BX356664
 KEYWORDS BX356664.1 GI.30378083
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0D1015CA02NP1.

FEATURES

source

1.1201
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1015YB03"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.5%; Score 45; DB 13; Length 1201;
 Best Local Similarity 12.9%; Pred. No. 0.32; Indels 0; Gaps 0;
 Matches 47; Conservative 168; Mismatches 148; Indels 0; Gaps 0;
 QY 307 GTTTCGCGGCTGATTTTCGGGGCTCCACGAATGCTGTCATCGCCCTCGTTGTGG 366
 DB 745 STBST 804

QY 367 CGACGCGCCCTGCGGCGCATGCGCTCACCACTTGTGTGAACCTCCGCCATAT 426
 DB 805 SBTCTSSSSSSSSSBTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 864
 QY 427 CTATGCTTTTATCCCGCTGCATGTGTCAAAACCCATTCGCTTCTATTCGT 486
 DB 865 YSSBSST 924
 QY 487 TTTCGCGCTTATCGAAGACCTACGACTCATCTCGGCGAGCCCGCAGCTGTGCG 546
 DB 925 BTSSBSSTSTSTSSSSTSTSTSSSSSSSBTSKSSSSSSSSSSTSTSTSTTT 984
 QY 547 GTGCGCACTTATCTCAATGCAAAATAGCGTTCACTCTGATTCGCGGCTCAC 606
 DB 985 VTSSBSST 1044
 QY 607 CGAGTGCAGATCGACAGATGATTCCTTTGAATTAAGGCGCTGAGTCCCTTTG 666
 DB 1045 SSSBSSSSSSSSSSSSBTTTBTSTSTSTSTSTSTSTSTSTSTSTSTSTTT 1104
 QY 667 CTC 669
 DB 1105 TTS 1107

RESULT 3
 BX376097 1201 bp mRNA linear EST 08-MAY-2003
 LOCUS BX376097 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC022YM12 5-PRIME, mRNA sequence.
 ACCESSION BX376097
 VERSION BX376097
 KEYWORDS BX376097.1 GI.30434756
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 2866.f
 Contact: Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DC022B06P1.

FEATURES

source

1.1201
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC022YM12"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.3%; Score 41.8; DB 13; Length 1201;
 Best Local Similarity 9.1%; Pred. No. 2.7;
 Matches 55; Conservative 238; Mismatches 310; Indels 2; Gaps 1;
 QY 593 TTTCGCGCTTCACCGAGTGCAGATGCTGATTCCTTTGAATTAAGGCTTC 652
 DB 441 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 500
 QY 653 GAGTTGCGCTTGTCTCTCTTGTCAAGCTGACTTGATTCCTGCGCAAGAAAAAG 712

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 415.78 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6
Perfect score: 1271
Sequence: 1 ggcggatcaatggaatctgag.....agccttaagcagacagcgc 1271

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_huv:*
19: em_gse_pin:*
20: em_gse_vtc:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vtl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.6	4.3	473	28	AO936633 HSJ41-905
2	45	3.5	1201	13	BX356664 BX356664
3	41.8	3.3	1201	13	BX376097 BX376097
4	40.8	3.2	1201	13	BX381961 BX381961

C 5	39.4	3.1	692	13	BX540413	BX540413
C 6	39.4	3.1	718	9	AJ397466	AJ397466
C 7	39.4	3.1	743	29	CC684922	CC684922
C 8	39.4	3.1	756	13	BU446643	BU446643
C 9	39.4	3.1	787	29	CNS009WA	ALU54013 Drosoph11
C 10	39.4	3.1	885	13	BX425603	ALU54013 Drosoph11
C 11	39.4	3.1	885	13	CNS005MG	ALU60428 Drosoph11
C 12	39.4	3.1	859	13	BF341287	BF341287 602013322
C 13	39.4	3.1	859	13	CNS0091P	ALU53013 Drosoph11
C 14	38.6	3.0	925	29	CNS0091P	BX361080
C 15	38.6	3.0	1201	13	BX361080	BX361080
C 16	38.6	3.0	835	13	CB57749	CB57749 OSJTEC3F
C 17	37.6	2.9	997	29	CNS005TE	ALU60767 Drosoph11
C 18	37.4	2.9	521	9	AJ449936	AJ449936
C 19	37.4	2.9	697	28	BH887860	BH887860 LB01712a
C 20	37.4	2.9	1108	13	BX382084	BX382084
C 21	37.2	2.9	423	28	B44402	ALU513871
C 22	37.2	2.9	730	12	BG845919	B44402 HS-1059-A2
C 23	37.2	2.9	758	29	CG793422	BG845919 1024011FO
C 24	37.2	2.9	1454	10	BF101158	CG793422 ZMABD030
C 25	37.2	2.9	535	9	A1512075	BF101158 601754733
C 26	37.2	2.9	603	9	A1258415	A1512075 LD43759.5
C 27	37.2	2.9	604	9	A1261160	A1258415 LP01614.5
C 28	37.2	2.9	608	9	A1514464	A1261160 LP01614.5
C 29	37.2	2.9	613	9	A1258135	A1514464 LD40978.5
C 30	37.2	2.9	616	12	B1641382	A1258135 GH14135.5
C 31	37.2	2.9	681	12	B1586164	B1641382 SD24651.5
C 32	36.8	2.9	464	28	BH805539	B1586164 RH26015.5
C 33	36.8	2.9	649	12	BG868786	BH805539 1008061A1
C 34	36.8	2.9	1201	29	CNS016BR	BG868786 602787626
C 35	36.6	2.9	539	29	BE013856	ALU06545 Drosoph11
C 36	36.6	2.9	542	29	CNS00CT2	BE013856 125354 MA
C 37	36.6	2.9	1134	13	BX395008	ALU059804 Drosoph11
C 38	36.4	2.9	288	12	CC325053	BX395008 EX395008
C 39	36.4	2.9	910	28	CC362201	BX325053
C 40	36.2	2.8	673	10	BJ287010	CC362201 PUHK887D
C 41	36.2	2.8	827	29	CC669295	BJ287010
C 42	36.2	2.8	837	13	BX331708	CC669295 OGPA017TH
C 43	36.2	2.8	902	13	CG358991	CG629295 OGPA017TH
C 44	36.2	2.8	962	29	BQ350424	BX331708
C 45	36.2	2.8	1101	13	BQ678281	CG358991 OG3A135TH
						BQ350424 AGENCOURT
						BQ678281 AGENCOURT

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PIRME
COMMENT

AO936633
HSJ41-905 Human NotI clones Homo sapiens genomic, genomic survey
sequence.
AO936633
AO936633.1 GI:7213011
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Zabarovskiy,E.R., Gizatuljin,R., Podowski,R.M., Zabarovskaya,V.V.,
Xie,L., Muravenko,O.V., Kozhev,S., Petrenko,L., Skobeleva,N.,
Li,J., Protodopov,A., Kashuba,V., Erbergh,I., Winberg,G. and
Wahlstedt,C.
NotI clones in the analysis of the human genome
Nucleic Acids Res. 28 (7), 1635-1639 (2000)
10710430
Contact: Podowski RM
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337983

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rnpn

Page 6

[illegible]

Search completed: April 17, 2004, 20:13:02
Job time : 148.874 secs

RESULT 12
US-10-021-698A-716
Sequence 716, Application US/10021698A

GENERAL INFORMATION:

APPLICANT: KEITH, TIM

APPLICANT: LITTLE, RANDALL

APPLICANT: VAN EERDENEGH, PAUL

APPLICANT: DUPUIS, JOSEF

APPLICANT: DEL MASTRO, RICHARD

APPLICANT: SIMON, JASON

APPLICANT: ALLEN, KRISTINA

APPLICANT: PANDIT, SUNIL

TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO

TITLE OF INVENTION: RESPIRATORY DISEASES AND OBESITY

FILE REFERENCE: 29/6-404US1

CURRENT APPLICATION NUMBER: US/10/021,698A

CURRENT FILING DATE: 2001-10-22

PRIOR APPLICATION NUMBER: 60/211,749

PRIOR FILING DATE: 2000-06-14

NUMBER OF SEQ ID NOS: 6160

SOFTWARE: Patentin 2.1

SEQ ID NO 716

LENGTH: 173233

TYPE: DNA

ORGANISM: Homo sapiens

NAME/KEY: modified_base

LOCATION: (88987)..(89086)

OTHER INFORMATION: a, t, c or g

FEATURE:

NAME/KEY: modified_base

LOCATION: (167286)..(167385)

OTHER INFORMATION: a, t, c or g

US-10-021-698A-716

Query Match 2.6%; Score 33.2; DB 6; Length 173233;

Best Local Similarity 55.1%; Pred. No. 39;

Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1052 AGACGTACCTTGTGCTCATTCGCGTGGATTAAGTGGGCGCATCTTCTGCG 1111

DB 84488 AGATTAAGCTTGAACCAATATGTTACTTATCTTTGTTGTTGTTCTCG 84547

QY 1112 GTGAGCAGCACTTGTGAGCGTGGCGCTGGACCATCGTTTGTGACTGTGAA 1169

DB 84548 GTAACTCAACCTGTGCTGTTGGCTGTCACTGATGATTTGACATGAA 84605

RESULT 13
US-60-548-091-19934
Sequence 19934, Application US/60548091

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001506

CURRENT APPLICATION NUMBER: US/60/548,091

CURRENT FILING DATE: 2004-02-27

NUMBER OF SEQ ID NOS: 24433

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19934

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-60-548-091-19934

Query Match 2.6%; Score 33; DB 7; Length 201;

Best Local Similarity 52.7%; Pred. No. 0.63;

Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 342 TGCTGTCATCGCCCTGCTGTTGTGGGCGCAGCCCTCGGCGCATCGGCTCACCACAT 401

DB 106 GCTTCAGAGCAGCCAGCAATTTCACTCACTGAGCGTCCGCTGCGGATTTAC 165

QY 462 ACCCATTTGCC 472

DB 166 TGCCCTTGCC 176

RESULT 14
US-60-548-091-19938
Sequence 19938, Application US/60548091

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele et al.

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001506

CURRENT APPLICATION NUMBER: US/60/548,091

CURRENT FILING DATE: 2004-02-27

NUMBER OF SEQ ID NOS: 24433

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 19938

LENGTH: 201

TYPE: DNA

ORGANISM: Homo sapiens

US-60-548-091-19938

Query Match 2.6%; Score 33; DB 7; Length 201;

Best Local Similarity 52.7%; Pred. No. 0.63;

Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 342 TGCTGTCATCGCCCTGCTGTTGTGGGCGCAGCCCTCGGCGCATCGGCTCACCACAT 401

DB 49 TGCTGTCATCGCCCTGCTGTTGTGGGCGCAGCCCTCGGCGCATCGGCTCACCACAT 108

QY 402 TGCTGTCATCGCCCTGCTGTTGTGGGCGCAGCCCTCGGCGCATCGGCTCACCACAT 461

DB 109 GCTTCAGAGCAGCCAGCAATTTCACTCACTGAGCGTCCGCTGCGGATTTAC 168

QY 462 ACCCATTTGCC 472

DB 169 TGCCCTTGCC 179

RESULT 15
US-10-767-701-9054
Sequence 9054, Application US/10767701

GENERAL INFORMATION:

APPLICANT: KOVALIC, David K.

APPLICANT: Zhou, Yihua

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(5335)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9054

LENGTH: 1028

TYPE: DNA

ORGANISM: Sorghum bicolor

OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1

US-10-767-701-9054

Query Match 2.6%; Score 32.8; DB 6; Length 1028;

Best Local Similarity 54.0%; Pred. No. 2.1;

Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 709 AAAGCAGATCCCTCTGCTGCTGCGAGTTTGAGGTTGACATGCTTGTGTAAT 768

Query Match 2.7%; Score 34; DB 6; Length 561;
Best Local Similarity 54.9%; Pred. No. 0.56; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 302 CCACGTGTTTCGGGCTGATTTTCGGGGCTCCACGAAATGCTGATCGCCCTGTT 361
DB 333 CCACGAGCGGCGACCTTGAGCTTCTGACCGGAGCGCGATCTCGGCTGCTCTGTC 274
QY 362 GTGGGGCGACCGCCCTGGGGCCCATCGGCTCACCAATTCGTGTAACTTCGGCCAC 421
DB 273 ACCAGGCGGACGCGCTTGCGGCTTCTCTCCACGACCTTGATGATGAGTACGCGCC 214
QY 422 GT 423
DB 213 CT 212

RESULT 9
US-10-765-790-62
Sequence 62; Application US/10765790
GENERAL INFORMATION:
APPLICANT: Bayer Healthcare LLC
APPLICANT: Beard, Chris
APPLICANT: Burgess, Chris
APPLICANT: Gannon, Allison
APPLICANT: Harvey, Jeanne
APPLICANT: Lechner, John F.
APPLICANT: Li, Zheng
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
FILE REFERENCE: 1657/2035
CURRENT FILING DATE: 2004-01-27
PRIOR FILING DATE: 2003-12-16
PRIOR APPLICATION NUMBER: US 10/737,082
NUMBER OF SEQ ID NOS: 300
SOFTWARE: Patent Invention 3.2
SEQ ID NO 62
LENGTH: 23490
TYPE: DNA
ORGANISM: Homo sapiens
US-10-765-790-62

Query Match 2.7%; Score 33.8; DB 6; Length 23490;
Best Local Similarity 49.7%; Pred. No. 7;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 301 CCCACTGTTTCCGGCTGATTTTCGGGGCTCCACGAAATGCTGATCGCCCTGTT 360
DB 21210 CTGGGTCTCTCCCTCCCTGATCTCTGTTCCACCTTCTGCTCTGAGGCTTGGA 21269
QY 361 TGTGGGGCGACCGCCCTGGGGCCCATCGGCTCACCAATTCGTGTAACTTCGGCA 420
DB 21270 GAGAGTGGGGCGGGGCTGGGCTGAGCTGAGCGGCTATGCGCTCTGAGGCGCTGACT 21329
QY 421 CGATTCTATGCGTTTTCATTCCTCCGCTGATGTGTCAAAACCCATTGCC 473
DB 21330 CTTCTCATGCTCTTCTTCTGCTGTGAGGGGGCGGTATACCACTGCGCC 21382

RESULT 10
US-60-548-091-5686
Sequence 5686; Application US/60548091
GENERAL INFORMATION:
APPLICANT: CARBELL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001506
CURRENT FILING DATE: US/60/548,091
NUMBER OF SEQ ID NOS: 24433
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5686

LENGTH: 24477
TYPE: DNA
ORGANISM: Homo sapiens
US-60-548-091-5686

Query Match 2.7%; Score 33.8; DB 7; Length 24477;
Best Local Similarity 51.9%; Pred. No. 7.1;
Matches 68; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 342 TGGTGTATCGCGCTGTTGTGGGGCGACGCGCCCTGGGGCGGCTGATCGGACACAT 401
DB 17603 TGGTGTATCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 17662
QY 402 TGGTGTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
DB 17663 GCTTCAGAGACACCGGACATTTCAATTCATGCGCTTCCGCTGCTGCTGCTGCTGCTGCT 17722
QY 462 ACCCATTTGCC 472
DB 17723 TGCCCTGCGC 17733

RESULT 11
US-09-804-291A-500
Sequence 500; Application US/09804291A
GENERAL INFORMATION:
APPLICANT: ZOZULA, SERGEY
TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 10037.54287US
CURRENT FILING DATE: US/09/804,291A
CURRENT FILING DATE: 2001-03-13
PRIOR FILING DATE: 2000-03-13
PRIOR FILING DATE: 2000-03-13
PRIOR FILING DATE: 2000-03-24
PRIOR FILING DATE: 2000-03-24
PRIOR FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-04-24
PRIOR FILING DATE: 2000-04-24
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-06-23
PRIOR FILING DATE: 2000-06-23
PRIOR FILING DATE: 2000-08-16
PRIOR FILING DATE: 2000-08-16
PRIOR FILING DATE: 2000-09-07
PRIOR FILING DATE: 2000-09-07
PRIOR FILING DATE: 2001-02-07
PRIOR FILING DATE: 2001-02-07
NUMBER OF SEQ ID NOS: 531
SOFTWARE: Patent Invention 3.2
SEQ ID NO 500
LENGTH: 930
TYPE: DNA
ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match 2.6%; Score 33.2; DB 5; Length 930;
Best Local Similarity 54.0%; Pred. No. 1.4;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 567 AATATGCTTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
DB 509 ACATTTCAATCACTATCTGTACTATGAGGCTCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 568
QY 627 TGATTCCTTTGAATTAAGGCTCGAGTTGCGCTTGTCTCTCTTTGTCAGGCTGA 686
DB 569 CAGGCTCTTGAAGTATGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628
QY 687 CTTTGG 692
DB 629 TGCTGG 634

```
; Sequence 983, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIORITY FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-983

Query Match      2.7%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 1;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 344 CTGTCATGCGCCCTGTGTGGCGGAGCGCCCTGTGGCGGCATGCGCTCACACA 400
Db 2360 CAGCTCCCGCCCGCAGTGTGGCTCAGGGCTCTGGCGCCATGCGTCTACACA 2304

RESULT 6
US-10-100-683-11511
; Sequence 11511, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIORITY FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-983
```

```
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11511
; LENGTH: 7071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-11511

Query Match      2.7%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.7;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 344 CTGTCATGCGCCCTGTGTGGCGGAGCGCCCTGTGGCGGCATGCGCTCACACA 400
Db 785 CAGCTCCCGCCCGCAGTGTGGCTCAGGGCTCTGGCGCCATGAGTCTCACACA 841

RESULT 7
US-10-767-701-7545
; Sequence 7545, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIORITY FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 7545
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
US-10-767-701-7545

Query Match      2.7%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.46;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Cy 315 GCTGATTTTGGGGGCTCCAGGAATGCTGTCATGCCCTCGTTGTGGGGCGAGCGC 374
Db 31 GCCGACACACACAGGGGAGAGACCCGCTCATCAACCGTTCGTGGAGCGGCGCGC 90

Cy 375 CCTGGGCGCCATGCGGCTCACACCAATGCTGTGAATTCGCCACAG 422
Db 91 CGCTGGCGTCCCTGGCGCTGACGCCGCGTCTGTGTGCGTGGCGAG 138

RESULT 8
US-10-767-701-744/c
; Sequence 744, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIORITY FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 744
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
US-10-767-701-744
```

```

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/417,884A
FILING DATE: 17-Apr-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051,571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEetical: NO
ORIGINAL SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-10-417-884A-1266
Query Match
Best Local Similarity 49.3%; Score 39.8; DB 6; Length 768;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 207 TAAAAACCTCCCTGCTGCGAGTGGGCGATGACCCGATGGTATGGCTTTGCTCT 266
DB 92 TCAAGACAGGTTACCTACCGTTTCGGTTATATCGGATTTGACTGATTTGGATCG 151
QY 267 TGGTATTCAATACGCTACGAATGGGCGAGCCCACTGTTTCGGGCTGATTTTCG 326
DB 152 TTGGGAAGAGTGGCGGATTTCAATCCACTAGTGTCAAGTTGATGTCCTGGCTGTAG 211
QY 327 GGGGCTCACCGAAATCTGTGATGCGCTGTTGTGGGCGAGGCGCCCTGGGCGCA 386
DB 212 CTGGTTCTGCCAATATATACAGTCAAGTGGCTTCTATGATCAAGCCCAATGCTTTCA 271
QY 387 TCGGCTCACCACTGCTGTGTAATTCG 417
DB 272 TCGTTTTCGACCTTTCTAGTCAATTCG 302

```

```

RESULT 3
US-10-767-701-5263
Sequence 5263, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.

```

```

APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21 (53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5263
LENGTH: 619
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS111974_1
US-10-767-701-5263

```

```

Query Match
Best Local Similarity 54.1%; Score 35.4; DB 6; Length 619;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 326 GCGGCTCCACCGAATGCTGATCGCCCTGCTGGGCGCGAGCCCTGGGCGCC 385
DB 132 GCGGCTTCACTTCTGCTGCTGCTGCGGCGCTGCGGCGAGCGGCAAGTGTGCGCTC 191
QY 386 ATGCGCTCACACATTTGCTGGAATTCGCGCAGATTTCTATGCTTTTATCCG 445
DB 192 ATCTGCGGAGCGCGCTGCTGATCTTCAAGCCCGTGTGCGCGCGGCGCATCAG 251
QY 446 CTGCATGCTGCA 458
DB 252 GTGGCGCTGCTCA 264

```

```

RESULT 4
US-10-767-701-10866
Sequence 10866, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21 (53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 10866
LENGTH: 1034
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
US-10-767-701-10866

```

```

Query Match
Best Local Similarity 49.7%; Score 34.6; DB 6; Length 1034;
Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 341 ATGCTGCTATGCGCTGCTGCTGCGAGCGCCCTGGGCGCGCATGCGCTCACACA 400
DB 151 ATGCGGCGGCTTCTATGAGGCGAGGCGCTGCGCGCACCGGCGGCGCGCGCCCT 210
QY 401 TTGCTGTGAATCTCCGCGCATTTATGCGCTTTTATTCCTCCGCTGATGTGTA 460
DB 211 GTGGGCGGAGTGTCCCGCGAGTTTTCGCGCGCGCTACGTTGTTGCTTACGTAAG 270
QY 461 AACCCATTTGCGGCTTCTATGCTTTCGCTGCTTTCGCTTTCGAGCAAGCTACAGTC 517
DB 271 AAGGCTACAGCTTCTTCGAGCGGCGCATTCACGCTACCGAGCGCAAGGCGCGCTC 327

```

```

RESULT 5
US-10-100-683-983/c

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 146.874 Seconds
(without alignments)
5694.647 Million cell updates/sec

Title: US-10-608-504-6
Perfect score: 1271
Sequence: 1 gcggcagcaacgaagatctag.....agccttaagcagacagcgc 1271

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1362413 seqs, 329030602 residues

Total number of hits satisfying chosen parameters: 2724826

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New:*
1: /cgn2_6/ptodata/1/pna/PCF_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.2	3.2	1014	6 US-10-767-701-10854	Sequence 10854, A
2	39.8	3.1	768	6 US-10-417-884A-1266	Sequence 1266, Ap
3	35.4	2.8	619	6 US-10-767-701-5263	Sequence 5263, Ap
4	34.6	2.7	1034	6 US-10-767-701-10866	Sequence 10866, A
5	34.6	2.7	3173	6 US-10-683-983	Sequence 983, App
6	34.6	2.7	7071	6 US-10-100-683-11511	Sequence 11511, A
7	34.4	2.7	666	6 US-10-767-701-7545	Sequence 7545, Ap
8	34	2.7	561	6 US-10-767-701-744	Sequence 744, App
9	33.8	2.7	23490	6 US-10-765-790-62	Sequence 62, App1
10	33.8	2.7	24477	7 US-60-548-091-5686	Sequence 5686, Ap
11	33.2	2.6	930	6 US-09-804-291A-500	Sequence 500, App
12	33.2	2.6	930	6 US-10-021-698A-716	Sequence 716, App
13	33	2.6	201	7 US-60-548-091-19934	Sequence 19934, A
14	33	2.6	201	7 US-60-548-091-19934	Sequence 19934, A
15	32.8	2.6	1028	6 US-10-767-701-9054	Sequence 9054, Ap
16	32.6	2.6	519	6 US-10-767-701-23189	Sequence 23189, A
17	32.6	2.6	1082	6 US-10-767-701-14711	Sequence 14711, A
18	32.6	2.6	438429	6 US-10-796-307-8767	Sequence 8767, Ap
19	32.4	2.5	1044	6 US-10-775-920-132	Sequence 132, App
20	32.4	2.5	1076	6 US-10-775-920-136	Sequence 136, App
21	32.4	2.5	1273	7 US-60-552-390-626	Sequence 626, App
22	32.4	2.5	1273	7 US-60-552-390-626	Sequence 626, App
23	32.4	2.5	1278	6 US-10-775-920-130	Sequence 130, App
24	32.4	2.5	1292	7 US-60-552-390-624	Sequence 624, App
25	32.4	2.5	1292	7 US-60-552-390-624	Sequence 624, App
26	32.4	2.5	1297	6 US-10-775-920-131	Sequence 131, App

27	32.4	2.5	1297	6 US-10-775-920-133	Sequence 133, App
28	32.4	2.5	1306	6 US-10-775-920-134	Sequence 134, App
29	32.4	2.5	1465	7 US-60-552-390-623	Sequence 623, App
30	32.4	2.5	1465	7 US-60-552-390-623	Sequence 623, App
31	32.4	2.5	1532	7 US-60-552-390-627	Sequence 627, App
32	32.4	2.5	1532	7 US-60-552-390-627	Sequence 627, App
33	32.4	2.5	1625	7 US-60-552-390-625	Sequence 625, App
34	32.4	2.5	1625	7 US-60-552-390-625	Sequence 625, App
35	32.4	2.5	1625	7 US-60-552-390-625	Sequence 625, App
36	32.2	2.5	228654	7 US-60-548-091-5633	Sequence 5633, Ap
37	32.2	2.5	321019	6 US-10-803-180-1684	Sequence 1684, Ap
38	32.2	2.5	321019	7 US-60-548-091-5604	Sequence 5604, App
39	32	2.5	1202	6 US-10-796-307-259	Sequence 259, App
40	32	2.5	1271	6 US-10-796-307-259	Sequence 259, App
41	32	2.5	32644	6 US-10-796-307-8771	Sequence 8771, Ap
42	32	2.5	84156	6 US-10-796-307-8770	Sequence 8770, Ap
43	31.8	2.5	112486	6 US-10-767-471-10642	Sequence 10642, A
44	31.6	2.5	715	6 US-10-767-701-3886	Sequence 3886, Ap
45	31.6	2.5	5448	1 PCT-US03-38685-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-10-767-701-10854
Sequence 10854, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kowalc, David K.
APPLICANT: Zhou, Yihua
TITLE OR INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ. ID NOS: 63128
SEQ. ID NO 10854
LENGTH: 1014
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS47627_1
US-10-767-701-10854

Query Match 3.2%; Score 40.2; DB 6; Length 1014;
Best Local Similarity 49.3%; Pred. No. 0.0066;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 231 TGGGATGTACCCGATGTATGCGTTGCTGTTATTCAATAGGCTACGAAT 290
DB 22 TGGGATGTACCCGATGTATGCGTTGCTGTTATTCAATAGGCTACGAAT 81
QY 291 GGTGGAGAGCCCACTGTTTCCGGCTGATTTCCGGGCTTCACCGAATCTGTCTA 350
DB 82 AGTCTGGGGCCCTCTCTTGGCAAGGTCTTCCGGGCTTCACCGAATCTGTCTA 141
QY 351 TCGCCCTGCTTGGGGGAGAGCCCTCGGGCCCATTCGGGCTTCACCGAATCTGTCTA 410
DB 142 GCCGCTCTGAGGGCTTGGAGCCCTCGGGCCCATTCGGGCTTCACCGAATCTGTCTA 201
QY 411 ACTTCGCGCAATTCATGCTGTTTCATTC 443
DB 202 GATCCGCTCTGAGGGCTTGGAGCCCTCGGGCCCATTCGGGCTTCACCGAATCTGTCTA 234

RESULT 2
US-10-417-884A-1266
Sequence 1266, Application US/10417884A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OR INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rmpb

Page 10

PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6067
LENGTH: 594
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURES:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-10-156-761-6067

Query Match 3.4%; Score 43; DB 14; Length 594;
Best Local Similarity 50.7%; Pred. No. 0.0029;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCAGGAGCTCCACCAAAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 383
DB 203 TCACGGCTGCGCCATCGCGAGGCTGCGCATGCTGCGACCGCGCTCGGGTGG 262
QY 384 CCATCGCGCTCACCAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 443
DB 263 GCAAGTCCGACGACGATCGCGGATCGTCTGCTTCTTCTGCGGCTACTGCTCA 322
QY 444 CGCTGATGCTGTCACAAAACCCGATGCGCCCTTCTATTCGCTTTCGCGCTTATCGACG 503
DB 323 CGTTCGCGGAGTGTGTAAGGCGGCTGTCGCGCTTCGACCGCCCTGCGGGTGGCGCTCG 382
QY 504 AAGCTACGCACTCACTGGGCGC 526
DB 383 CGCGGACACCTGTCATCGCC 405

RESULT 14

US-10-156-761-1/C
Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURES:
NAME/KEY: misc_feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.4%; Score 43; DB 14; Length 9025608;
Best Local Similarity 50.7%; Pred. No. 1.4;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 324 TCAGGAGCTCCACCAAAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 383
DB 7327430 TCACGGCTGCGCCATCGCGAGGCTGCTGCGCATGCTGCGACCGCGCTCGGCTGG 7327371

QY 384 CCATGCGCTCACCAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 443
DB 7327370 GCAAGTCCGACGACGATCGCGGATCGTCTGCTTCTTCTGCGGCTACTGCTCA 7327311
QY 444 CGCTGATGCTGTCACAAAACCCGATGCGCCCTTCTATTCGCTTTCGCGCTTATCGACG 503
DB 7327310 CGTTCGCGGAGTGTGTAAGGCGGCTGTCGCGCTTCGACCGCCCTGCGGGTGGCGCTCG 7327251
QY 504 AAGCTACGCACTCACTGGGCGC 526
DB 7327250 CGCGGACACCTGTCATCGCC 7327228

RESULT 15

US-10-184-644-80/C
Sequence 80, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zhenli
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
PRIOR FILING DATE: 2002-06-28
PRIOR APPLICATION removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 80
LENGTH: 351
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-80

Query Match

3.3%; Score 41.4; DB 14; Length 351;
Best Local Similarity 13.4%; Pred. No. 0.0073;
Matches 45; Conservative 103; Mismatches 187; Indels 0; Gaps 0;

QY 161 GAACGAGATGATTAAGGTTATCGCGCTACGAATCGCGGACGCTAATAAACCCTCCCTT 220
DB 350 GAD..B.Y.G.YGMB.GGMD.MSGV.T.B.TAD.CYCGG...GCH.TMSANNGAYCA 291
QY 221 GCTGAGGTTGGGACATGACCCGATGATGCGTTGCTCTTGGTATTCAATAC 280
DB 290 C.BCSYKH..G.H.S...RB.H.HAD.M..TNY.B.TYSSCBY..B.ATCH.M.M.TDC 231
QY 281 GGCTACGATGATGAGGCGCCGACCTGTTTCGCGGCTGATTTTCGCGGCTCACCGAA 340
DB 230 B.Y...BTYGANB.GSCSBDGNG.BYKDA..TMYCT.NND.NK.TYSSSTCMY.YKH 171
QY 341 ATGCTGATCGCCCTGTTGTTGGGCGACGCGCCCTGGGCGCATCGGCTCACCA 400
DB 170 B.S.S.TBCRT.NHSGSM..C..SBDABVHGGM...CSGTVAAT.TBCTSS.TRB. 111
QY 401 TTGCTGTAATTCGCGGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
DB 110 TT.TAC.ABMC.B..SH.ASGMYTY..R.G...T..TCYBSHKB..NYMK.B.Y. 51
QY 461 AACCCATGCGCGTTCTATTCGTTTCGCGCT 495
DB 50 ..CM.M.GA...S.SCSS.MT.R.MNSTTTS.BT 16

Search completed: April 17, 2004, 21:01:41
Job time : 642.884 secs

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rnpb

Page 8

	Matches	318;	Conservative	0;	Mismatches	296;	Indels	3;	Gaps	1
OY	222	CTGAGAGTTGGGCAATGTACCCGAAATTGGATAATTGGCGTTTGCNCTTGTTGTAATTCANPAG	201							
Db	53	CGGTCGGATTGGGAGCTGAATTCGCCGCTGGGTTTGAGCCTTTGGGCTGTATATGCTCACAGG	112							
OY	282	GCTACGAATGATGAGGACGCCCATCTGTTTTCCGAGCTGANTTTTGGCGGAGCTCCACGGAAA	341							
Db	113	GTTTCGCTGTGTGTGTGAGACGCCGATTTTCTCTTGATCTAATGCGGTTTCATGAAAT	172							
OY	342	TGCTGGATTCGCGCCTGTGTTGTGGGGGACGAGCCCCCTGGGGGCATGAGGCTACACAAAT	401							
Db	173	TTCGTGCAATCGGCAATGTCACCGACAGATATGGGCGCGTTTTCGGCGGGCGGTGCTGGTT	232							
OY	402	TGCTGTGAACCTTCGCCACGATATCTATGCGTTTCAATCCGCGTGCATGTGATCAAA	461							
Db	233	TCAATGGAAATTTCCGCCCAATTTTCTAAGGTCTACCTTCCACGCCAACCGCATCAAGT	292							
OY	462	ACCOCATGGCCGTTTCTTATTCGATTTTGGGCTTATGACAGAGCTTAGAGCTACTG	521							
Db	293	CCGGCGCGGCGCCCGCGCTTATTCACCTAAGCGCTTACCGAGAATCTTAAGCCATGTGT	352							
OY	522	CGGCCAGGCCCGACAGCTGTGTGGCGGTG--GCGATTATCTAATGCAAATAGCGTTTC	578							
Db	353	CAGCCGCCCCACCTGCGCATCATGATGAGACCGCGGAGCTTAACGTTCAAATTTTGTGCC	412							
OY	579	ACTGCTACTGGATTCGGGAGTCTCACCGAGATGGAGATGGACAATTCCTTTG	638							
Db	413	AAGCTCTGTGGGTATTCGCCAATATTTATGGCGCTTGATGTGTGCATAGTCTGCCAGT	472							
OY	639	AAATTAAGGCTTCGATTCGCCCTTTGCTCTCTCTTTGTCAAGCTGACTTTGAAATCT	698							
Db	473	ATCTAAAGGACATGATTTTGTGCCACCGCGCTGTTTGTGTGCTGCGCTGGAGAGCAT	532							
OY	699	GCGCAAGAAAAAGAGATCCCTCTGCTGCTGCTGCGAGGTTTGAAGTTCAACATTCATC	758							
Db	533	TCAAAATATACAGATATTTGTGCTGCCATTTATGCGAGTGTATGTGGCTGTGCTTCG	592							
OY	759	TGTGTGTAATTCAGGCTCAGGCGCTTATTTGCGGCGCTGCTGANTCTTGGATCTGTGA	818							
Db	593	GTTTTTGGGCGCCGACAGATGCTGTATGCTTTGACCAAGTACTTTTGTATCTCTC	652							
OY	819	CCATCCGATCTCTTC	835							
Db	653	TTCCTCGCGTCCGCTTC	669							
 RESULT 12 US-09-738-626-1/c Sequence 1, Application US/09738626 Publication No. US20020197605A1. GENERAL INFORMATION: APPLICANT: NAKAGAWA, SATOSHI APPLICANT: MIZOGUCHI, HIROSHI APPLICANT: ANDO, SEIKO APPLICANT: HAYASHI, MIKIRO APPLICANT: OCHIAI, KEIKO APPLICANT: YOKOI, HARUHIKO APPLICANT: TATEISHI, NAOKO APPLICANT: SENOH, AKIHITO APPLICANT: IKEDA, MASATO APPLICANT: OKAZKI, AKIO TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES FILE REFERENCE: 249-125 CURRENT APPLICATION NUMBER: US/09/738,626 CURRENT FILING DATE: 2000-12-18 PRIOR APPLICATION NUMBER: JP 99/377484 PRIOR FILING DATE: 1999-12-16 PRIOR APPLICATION NUMBER: JP 00/459162 PRIOR FILING DATE: 2000-04-07 PRIOR APPLICATION NUMBER: JP 00/280588 PRIOR FILING DATE: 2000-08-03 NUMBER OF SEQ ID NOS: 7059										

```

; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

```

Query Match	10.0%;	Score 127.4;	DB 9;	Length 3309400;
Best Local Similarity	51.5%;	Pred. No. 1.4e-29;		
Matches 318; Conservative	0;	Mismatches 296;	Indels 3;	Gaps 1

Oy	22	CNCAGCTTGGGACATGAACCCAAATTGTATTTGGTTTTGGCTCTTGATTTCAATAACG	281
Db	3289969	CGGTGGCTTTGGAGGCTGATTTCCGCTGGGTTTGGCTTTGGGCTGTGATGATGCAGACAG	3289910
Oy	282	GCTACGAATGTGGGACGCCCACTGTTTTCCGGCCTGATTTTCGGGGCTCCACCAGAA	341
Db	3289909	GTTTTCCCCTGGTGGTGAAGCCGCAATTTTCTTCCTTGATGATTAATGCGGGTTGATGAGAA	3289850
Oy	342	TGTGTGTAATGGCCCTGATTTGGAGGACAGGCCCTGGAGGCGCATTCGGGCTCACCAAT	401
Db	3289949	TTCTGGCAATCGGCAATGATCAACGCAAGTATCGGCCCCGTTTTGGGAGGCGTCCGTGTT	3289790
Oy	402	TGCTGTGAATTTCCGCAAGTATTTATGCGTTTTCAATTCGCGTGCAATGTGTCAAAA	461
Db	3289789	TCATGTGTAATTTCCGCAACAATTTCTACGCTCAACTTCCACGCCACGCATCAAGT	3289730
Oy	462	ACCCCAATGCCCGTTTCATTCGTTTTGGAGCTTATCGACGAAGCTAAGCGATCACTG	521
Db	3289729	CGGGCCCGCGCCGCGCCATTTCCACTCAACGGCTTACCAAGATCTTAGGCCATCGTGT	3289670
Oy	522	CGGCAAGGCCGCAAGCTGTGCGGCGTG---GCGATTAATCAATGCAAAATAGCTTC	578
Db	3289669	CAGCCCGCCCACTGCGCATATCAGTGGCAAGCGGGGTGCTAACCTCAAATTTGTGTC	3289610
Oy	579	ACTCTACAGGGTATTCGGCGGCTTCAAGCGGAATGGCGAATCCAGAATGATTCCTTTG	638
Db	3289609	AAGCTGTGGTTATCCCAAGAAATTTTGGCGCTTGGTTGCTCAAGTGCTGCCCGATG	3289550
Oy	639	AAATTAAGGGCTCGAGTTCGCCCTTTGCTCTCTCTTGTCAAGCTGAATTTGATTCCT	698
Db	3289549	ATCTAAAGGCTAGGATTTTCCCTGACCCGCGTGTTTGGTGTGGGATGGAGGAGCAT	3289490
Oy	699	GCGGAACGAAAAAGAGATCCCTCTCTCTGTGCTGCAGCAATTTAGACTTCAACATTCGCT	758
Db	3289489	TCAAAATATACAGGATTTATTCGTGCATTAATCCGCGTGTATTTGGCTCTGTTTTCCG	3289430
Oy	759	TTGTGTATATTCAGGTCAGGCTTATTTGGCGCGCTGTATCTTTGGGCTGTGTGA	818
Db	3289429	GTTTTGTGGCGCCGACAGATGCTGTATTCGCTTTGACCAAGTATTTTGAATCTTTC	3289370
Oy	819	CCATCGGATCTTCTTC	835
Db	3289369	TTCTCCGCTCGGCTTC	3289353
RESULT 13			
US-10-156-761-6067			
; Sequence 6067, Application US/10156761			
; Publication No. US20030119018A1			
GENERAL INFORMATION:			
APPLICANT: OMURA, SATOSHI			
APPLICANT: IKEDA, HARUO			
APPLICANT: ISHIKAWA, JUN			
APPLICANT: HORIKAWA, HIROSHI			
APPLICANT: SHIBA, TADAYOSHI			
APPLICANT: SAKAKI, YOSHIOKI			
APPLICANT: HATORI, MASAHIRA			
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
FILE REFERENCE: 249-262			
CURRENT APPLICATION NUMBER: US/10/156,761			
CURRENT FILING DATE: 2002-05-29			
PRIORITY APPLICATION NUMBER: JP 2001-204089			

APPLICANT: SAMM, HERMANN
 APPLICANT: EGGELING, LOHAR
 APPLICANT: PFEFFERLE, WALTER
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
 FILE REFERENCE: 7601/80525
 CURRENT APPLICATION NUMBER: US/10/608,504
 CURRENT FILING DATE: 2003-06-30
 PRIOR APPLICATION NUMBER: US 09/471,803
 PRIOR FILING DATE: 1999-12-23
 PRIOR APPLICATION NUMBER: DE 199 51 708.8
 PRIOR FILING DATE: 1999-10-27
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 4
 LENGTH: 324
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(324)
 OTHER INFORMATION: brn
 FEATURE:
 OTHER INFORMATION: ATCC14752
 US-10-608-504-4

Query Match 25.5%; Score 324; DB 15; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.3e-99;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 853 ATGACACTGATTTCTCCGTGATTTCTCTGTTGCGAGTATGCGATTAATTCTTT 912
 DB 1 ATGACACTGATTTCTCCGTGATTTCTCTGTTGCGAGTATGCGATTAATTCTTT 60
 QY 913 GCGCTCCGGCGGCGTTCCGTTCTTAATCCTTAAGCCCTTACGATCAATTGCGGC 972
 DB 61 GCGCTCCGGCGGCGTTCCGTTCTTAATCCTTAAGCCCTTACGATCAATTGCGGC 120
 QY 973 AAAATGGCATGTGATGCGACAGGAATCTTGCAATTTGACCGCATCAAGTTTGGC 1032
 DB 121 AAAATGGCATGTGATGCGACAGGAATCTTGCAATTTGACCGCATCAAGTTTGGC 180
 QY 1033 AGCAATGCGATGATCTGAAGACTTAACCTTTGCTTATGCGGTTGCCATTAAGTG 1092
 DB 181 AGCAATGCGATGATCTGAAGACTTAACCTTTGCTTATGCGGTTGCCATTAAGTG 240
 QY 1093 GTGGCGCATCTTTGCGCGGTGAGACGACCTTGTTGAGCGTTGGCGCTGGACCATCGT 1152
 DB 241 GTGGCGCATCTTTGCGCGGTGAGACGACCTTGTTGAGCGTTGGCGCTGGACCATCGT 300
 QY 1153 TTGTTGACTGCTGTAATCTTTTC 1176
 DB 301 TTGTTGACTGCTGTAATCTTTTC 324

RESULT 10
 US-10-134-640-1/c
 Sequence 1, Application US/10134640
 Publication No. US20030017555A1
 GENERAL INFORMATION:
 APPLICANT: Baehre, Brigitte
 APPLICANT: Kallinowski, Jörn
 APPLICANT: Puhler, Alfred
 APPLICANT: Mockel, Bettina
 APPLICANT: Pfeifferle, Walter
 TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
 FILE REFERENCE: 990109 BT
 CURRENT APPLICATION NUMBER: US/10/134,640
 CURRENT FILING DATE: 2002-04-30
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 1

LENGTH: 715
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (151)..(612)
 OTHER INFORMATION: lrp-Gen
 FEATURE:
 NAME/KEY: -10 signal
 LOCATION: (88)..(93)
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: -35 signal
 LOCATION: (62)..(67)
 OTHER INFORMATION:
 US-10-134-640-1

Query Match 14.6%; Score 185; DB 12; Length 715;
 Best Local Similarity 100.0%; Pred. No. 1.5e-51;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCATCATGATGATCTAGCTTCAATATATGCAACAATAGCTTGAAGTCCGAAC 60
 DB 185 GCGCATCATGATGATCTAGCTTCAATATATGCAACAATAGCTTGAAGTCCGAAC 126
 QY 61 TGGCAACAAACTACCCGCAATTTGTGTATGATTTGTAGTGTGCAAAAAACGAAGAT 120
 DB 125 TGGCAACAAACTACCCGCAATTTGTGTATGATTTGTAGTGTGCAAAAAACGAAGAT 66
 QY 121 TCATTCAAGCTGAGAGTGTGCGCATCAAGCAGCCCTGGAACAGATGATTAAGTTA 180
 DB 65 TCATTCAAGCTGAGAGTGTGCGCATCAAGCAGCCCTGGAACAGATGATTAAGTTA 6
 QY 181 TCGGC 185
 DB 5 TCGGC 1

RESULT 11
 US-09-738-626-3408
 Sequence 3408, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 3408
 LENGTH: 711
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3408

Query Match 10.0%; Score 127.4; DB 9; Length 711;
 Best Local Similarity 51.5%; Pred. No. 6.1e-32;

Db 181 TCGGGGCTACGAAATCGCGCAAGGCTTAAACCTCCCTTGTGAGGTTGGGATGTA 240
Qy 241 CCCGATTTGGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 300
Db 241 CCCGATTTGGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 300
Qy 301 CCCGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 360
Db 301 CCCGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 360
Qy 361 TGTGGGCGAGGCGCCCTGGGCGGCAATGGCGCTCACACATGTGTGAATCTTCGCGCA 420
Db 361 TGTGGGCGAGGCGCCCTGGGCGGCAATGGCGCTCACACATGTGTGAATCTTCGCGCA 420
Qy 421 CGTATTTCTATGGGTTTTCATTCGCGCTGAGTGTGCAAAAACCCATTGCCGTTCTA 480
Db 421 CGTATTTCTATGGGTTTTCATTCGCGCTGAGTGTGCAAAAACCCATTGCCGTTCTA 480
Qy 481 TTGGGTTTTCGCGCTTATGAGAGAACCTACGAGCTACGATGGGCGAGGCGGCGGCTG 540
Db 481 TTGGGTTTTCGCGCTTATGAGAGAACCTACGAGCTACGATGGGCGAGGCGGCGGCTG 540
Qy 541 GTGGGCGTGGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGTTATCGGCGG 600
Db 541 GTGGGCGTGGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGTTATCGGCGG 600
Qy 601 TCTCACCGAGATGGCGATGCGAGATGTAATCTCTTTTGAATTAAGGCGCTGAGTGGC 660
Db 601 TCTCACCGAGATGGCGATGCGAGATGTAATCTCTTTTGAATTAAGGCGCTGAGTGGC 660
Qy 661 CCTTGTCTCTCTTTGTGACGCTGACTTGTGATTCCTGCGCAAGAAACGAAATGCC 720
Db 661 CCTTGTCTCTCTTTGTGACGCTGACTTGTGATTCCTGCGCAAGAAACGAAATGCC 720
Qy 721 TTCTGTCTGTCTGCGAGGTTTGAAGCTTCAACATTCCTTGTGTGATTCAGTGTGAGC 780
Db 721 TTCTGTCTGTCTGCGAGGTTTGAAGCTTCAACATTCCTTGTGTGATTCAGTGTGAGC 780
Qy 781 CCTATTGGGCGGCTGTGATCTTCTGGGCTGTGATGCAATCCGCTACTTCTTCTGGG 840
Db 781 CCTATTGGGCGGCTGTGATCTTCTGGGCTGTGATGCAATCCGCTACTTCTTCTGGG 840
Qy 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGTGA 900
Db 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGTGA 900
Qy 901 GTGATTAATTTTGGGCTGCGGCGGCTGCGTCTTGAATCTTGAAGCGCTAGTGATGA 960
Db 901 GTGATTAATTTTGGGCTGCGGCGGCTGCGTCTTGAATCTTGAAGCGCTAGTGATGA 960
Qy 961 CAATTTGTGGGCAAAATGCGATGTGATGCGAGCAAGAACTCTTGCATTTTGAACGCA 1020
Db 961 CAATTTGTGGGCAAAATGCGATGTGATGCGAGCAAGAACTCTTGCATTTTGAACGCA 1020
Qy 1021 TCAACGTTTGGGCAATGCGATGTGATGCGAGCAAGAACTCTTGCATTTTGAACGCA 1080
Db 1021 TCAACGTTTGGGCAATGCGATGTGATGCGAGCAAGAACTCTTGCATTTTGAACGCA 1080
Qy 1081 GCGATTTACGATGTGAGGCACTTCTTGGGCTGTGAGCAAGAACTCTTGCATTTTGAACGCA 1140
Db 1081 GCGATTTACGATGTGAGGCACTTCTTGGGCTGTGAGCAAGAACTCTTGCATTTTGAACGCA 1140
Qy 1141 GCGACATGCTTTTGTGAGCTGTGATCTTTTCTAAATCTGCAATTAATTAACAAT 1200
Db 1141 GCGACATGCTTTTGTGAGCTGTGATCTTTTCTAAATCTGCAATTAATTAACAAT 1200
Qy 1201 CCGGATGCTCTCAATTTGAAGGGAATGCGGATTTTGAAGAACTGAGAAAGGCTTAAG 1260
Db 1201 CCGGATGCTCTCAATTTGAAGGGAATGCGGATTTTGAAGAACTGAGAAAGGCTTAAG 1260
Qy 1261 CAGACAGCGCT 1271

Db 1261 CAGACAGCGCT 1271

RESULT 3
US-09-738-626-1
Sequence 1, Application US/09738626
Publication No. US20020197605A1

GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIKOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: ISENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 100.0%; Score 1271; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGGATCAATGGAATGATCTAGCTTATATATGACAAATAGCTAGTGGGCGCAAC 60
Db 276729 GCGGATCAATGGAATGATCTAGCTTATATATGACAAATAGCTAGTGGGCGCAAC 276788
Qy 61 TGGCAACAAATCTACCGGCAATTTGTGTGATGTTAGTGTGCAAAAACGCAAGAT 120
Db 276789 TGGCAACAAATCTACCGGCAATTTGTGTGATGTTAGTGTGCAAAAACGCAAGAT 276848
Qy 121 TCATTCAAGCTGAGGTTGTGCGCATCCAGGCGAGCCCTGGAACAGATGATAAGGTTA 180
Db 276849 TCATTCAAGCTGAGGTTGTGCGCATCCAGGCGAGCCCTGGAACAGATGATAAGGTTA 276908
Qy 181 TCGGCGTACGAAATGCGGCAAGGCTTAAAAACCTCCCTGTCTGCAAGGTTTGGCATGTA 240
Db 276909 TCGGCGTACGAAATGCGGCAAGGCTTAAAAACCTCCCTGTCTGCAAGGTTTGGCATGTA 276968
Qy 241 CCCGATTTGGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 300
Db 276969 CCCGATTTGGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 277028
Qy 301 CCCGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 360
Db 277029 CCCGATTTGGCTTTGTCTCTGGTTATTCAATACGAGCTACGATGGTGGGACG 277088
Qy 361 TGTGGGCGAGGCGCCCTGGGCGGCAATGGCGCTCACACATGTGTGAATCTTCGCGCA 420
Db 277089 TGTGGGCGAGGCGCCCTGGGCGGCAATGGCGCTCACACATGTGTGAATCTTCGCGCA 277148
Qy 421 CGTATTTCTATGGGTTTTCATTCGCGCTGAGTGTGCAAAAACCCATTGCCGTTCTA 480
Db 277149 CGTATTTCTATGGGTTTTCATTCGCGCTGAGTGTGCAAAAACCCATTGCCGTTCTA 277208
Qy 481 TTGGGTTTTCGCGCTTATGAGAGAACCTACGAGCTACGATGGGCGAGGCGGCGGCTG 540

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCGATCAATGGAATCTAGCTTCAATATTTGCAAAATAGCTTGAAGGTGGGCAAC 60
Db 1 GGGCGATCAATGGAATCTAGCTTCAATATTTGCAAAATAGCTTGAAGGTGGGCAAC 60
QY 61 TGGCAACAAATACCCGGCAATTTGTGATGATTTGTAGTGGCAAAAAGCAAGAT 120
Db 61 TGGCAACAAATACCCGGCAATTTGTGATGATTTGTAGTGGCAAAAAGCAAGAT 120
QY 121 TCATTCAAGCCCTGAGAGTGTCCCATCCAAAGGCAAGCCCTGGAACCAATGATAAGGTTA 180
Db 121 TCATTCAAGCCCTGAGAGTGTCCCATCCAAAGGCAAGCCCTGGAACCAATGATAAGGTTA 180
QY 181 TGGGGGCTACGAATGCGGCAAGTCTAAACCTCCCTGTGAGGTTGGGCAATGTA 240
Db 181 TGGGGGCTACGAATGCGGCAAGTCTAAACCTCCCTGTGAGGTTGGGCAATGTA 240
QY 241 CCCGATTGATTCGCTTGTGCTCTTGTGTAATCAATACGCTACGAATGCTGGCAGC 300
Db 241 CCCGATTGATTCGCTTGTGCTCTTGTGTAATCAATACGCTACGAATGCTGGCAGC 300
QY 301 CCCGATTGATTCGCTTGTGCTCTTGTGTAATCAATACGCTACGAATGCTGGCAGC 360
Db 301 CCCGATTGATTCGCTTGTGCTCTTGTGTAATCAATACGCTACGAATGCTGGCAGC 360
QY 361 TGTGGGCGAGGCCCCCTGGGCGCATCGCGCTCAACACATGCTGTGAACCTTCGCA 420
Db 361 TGTGGGCGAGGCCCCCTGGGCGCATCGCGCTCAACACATGCTGTGAACCTTCGCA 420
QY 421 CGATTCTATGCTTTTCAATCCCGGTGATGAGTCAAAACCCCATGCCCCGTTCTA 480
Db 421 CGATTCTATGCTTTTCAATCCCGGTGATGAGTCAAAACCCCATGCCCCGTTCTA 480
QY 481 TTGGGTTTTCGCGCTTATCGACGAAAGCTCAAGCAGTCACTGGGCGGCGCGCAGCTG 540
Db 481 TTGGGTTTTCGCGCTTATCGACGAAAGCTCAAGCAGTCACTGGGCGGCGCGCAGCTG 540
QY 541 GTGGGCGTGGGCACTTATCTCAATGCAATAGGCTTCACTCTATGCGGTATTTGGCGG 600
Db 541 GTGGGCGTGGGCACTTATCTCAATGCAATAGGCTTCACTCTATGCGGTATTTGGCGG 600
QY 601 TCTCACCGGAGTGGCGATGCGACAGTGAATCTTTTGAATTAAGGCGCTGAGTTCCG 660
Db 601 TCTCACCGGAGTGGCGATGCGACAGTGAATCTTTTGAATTAAGGCGCTGAGTTCCG 660
QY 661 CCTTTGCTCTCTTTGTGACGCTGACTTGTGATTCCTGCGGACGAAAGCAATCCC 720
Db 661 CCTTTGCTCTCTTTGTGACGCTGACTTGTGATTCCTGCGGACGAAAGCAATCCC 720
QY 721 TTTCTGCTGCTGCGAGAGTTTGAAGTCAACATTTGCTTGTGATTCAGTCAAGG 780
Db 721 TTTCTGCTGCTGCGAGAGTTTGAAGTCAACATTTGCTTGTGATTCAGTCAAGG 780
QY 781 CCTATTGCGGCGCTGCTGATCTTTCTTGGGTCTGTGACATCCGTAATCTTCTTGGG 840
Db 781 CCTATTGCGGCGCTGCTGATCTTTCTTGGGTCTGTGACATCCGTAATCTTCTTGGG 840
QY 841 AAAGGCTGCTAAATGAACAAGATTTCTCGGTATTCCTGTTGTGCGAGATGGA 900
Db 841 AAAGGCTGCTAAATGAACAAGATTTCTCGGTATTCCTGTTGTGCGAGATGGA 900
QY 901 GTCAATTAATTTGGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGTGA 960
Db 901 GTCAATTAATTTGGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGTGA 960
QY 961 CAATTGTGGGCAAAATGGCGATGATGATGCGACAGGAATCCCTGCAATTTTGAACGCA 1020
Db 961 CAATTGTGGGCAAAATGGCGATGATGATGCGACAGGAATCCCTGCAATTTTGAACGCA 1020
QY 1021 TCAACGTTTGGGCAAAATGGCGATGATGATGCGACAGGAATCCCTGCAATTTTGGCGGT 1080
Db 1021 TCAACGTTTGGGCAAAATGGCGATGATGATGCGACAGGAATCCCTGCAATTTTGGCGGT 1080

Db 1021 TCAACGTTTGGGCAAAATGGCGATGATGATGCGACAGGAATCCCTGCAATTTTGGCGGT 1080
QY 1081 GCGATTACAGATGATGCGGATCTTCTTGGCGGTCGACGACCTTGTGAGCGTGGCGCT 1140
Db 1081 GCGATTACAGATGATGCGGATCTTCTTGGCGGTCGACGACCTTGTGAGCGTGGCGCT 1140
QY 1141 GCGACCATGCTTTTGTGATGATGATCTTTTAAACCTGCAATAAATAAATAAT 1200
Db 1141 GCGACCATGCTTTTGTGATGATGATCTTTTAAACCTGCAATAAATAAATAAATAAT 1200
QY 1201 CCGGATGCCCTCAATTTGAAGGGGATCGGATTTTAAAGAACCTAGAAAAGGCTTAAG 1260
Db 1201 CCGGATGCCCTCAATTTGAAGGGGATCGGATTTTAAAGAACCTAGAAAAGGCTTAAG 1260
QY 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

RESULT 2

US-10-608-504-6
Sequence 6, Application US/10608504
Publication No. US20040014123A1
GENERAL INFORMATION:
APPLICANT: KENNERNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGELING, LOTHAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
FILE REFERENCE: 7601/80525
CURRENT APPLICATION NUMBER: US/10/608,504
PRIOR APPLICATION NUMBER: US 09/471,803
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnf
FEATURE:
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnf
FEATURE:
OTHER INFORMATION: ATCC33032
US-10-608-504-6

Query Match 100.0%; Score 1271; DB 15; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCGATCAATGGAATCTAGCTTCAATATTTGCAAAATAGCTTGAAGGTGGGCAAC 60
Db 1 GGGCGATCAATGGAATCTAGCTTCAATATTTGCAAAATAGCTTGAAGGTGGGCAAC 60
QY 61 TGGCAACAAATACCCGGCAATTTGTGATGATTTGTAGTGGCAAAAAGCAAGAT 120
Db 61 TGGCAACAAATACCCGGCAATTTGTGATGATTTGTAGTGGCAAAAAGCAAGAT 120
QY 121 TCATTCAAGCCCTGAGAGTGTCCCATCCAAAGGCAAGCCCTGGAACCAATGATAAGGTTA 180
Db 121 TCATTCAAGCCCTGAGAGTGTCCCATCCAAAGGCAAGCCCTGGAACCAATGATAAGGTTA 180
QY 181 TGGGGGCTACGAATGCGGCAAGTCTAAACCTCCCTGTGAGGTTGGGCAATGTA 240
Db 181 TGGGGGCTACGAATGCGGCAAGTCTAAACCTCCCTGTGAGGTTGGGCAATGTA 240

us-10-608-504-6.rnpk

GenCore version 5.1.6
(c) 1993 - 2004 Compu

Search time 617.884 Seconds
(without alignment)

9204.305 Million cell updates/sec

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

assidues

meters: 5780264

U *

[illegible]

redicted by chance to have a score of the result being printed, total score distribution.

150

[illegible]

C 16	41.4	3.3	331	14	US-10-184-633-80	Sequence 80, App
C 17	40	3.1	671	14	US-10-184-634-346	Sequence 346, App
C 18	40	3.1	671	14	US-10-184-634-346	Sequence 346, App
C 19	39.8	3.1	556	14	US-10-184-634-310	Sequence 310, App
C 20	39.8	3.1	556	14	US-10-184-634-310	Sequence 310, App
C 21	39.8	3.1	556	14	US-10-063-685-100	Sequence 100, App
C 22	38.4	3.0	594	12	US-10-142-426-10	Sequence 10, App
C 23	38.4	3.0	594	14	US-10-123-155-10	Sequence 10, App
C 24	38.4	3.0	594	14	US-10-146-731-10	Sequence 10, App
C 25	38.4	3.0	594	14	US-10-140-447-10	Sequence 10, App
C 26	38.4	3.0	594	14	US-10-141-761-10	Sequence 10, App
C 27	38.4	3.0	594	14	US-10-142-885-10	Sequence 10, App
C 28	38.4	3.0	594	14	US-10-158-790-10	Sequence 10, App
C 29	38.4	3.0	594	15	US-10-137-871-10	Sequence 10, App
C 30	38.4	3.0	594	15	US-10-140-923-10	Sequence 10, App
C 31	38.4	3.0	594	15	US-10-141-756-10	Sequence 10, App
C 32	38.4	3.0	594	15	US-10-141-759-10	Sequence 10, App
C 33	38.4	3.0	594	15	US-10-140-805-10	Sequence 10, App
C 34	38.4	3.0	594	15	US-10-140-864-10	Sequence 10, App
C 35	38.2	3.0	946	9	US-09-778-320-11	Sequence 211, App
C 36	38.2	3.0	946	9	US-09-910-689-11	Sequence 211, App
C 37	38.2	3.0	946	13	US-10-010-742-21	Sequence 211, App
C 38	37.8	3.0	457	14	US-10-184-644-48	Sequence 48, App
C 39	37.8	3.0	457	14	US-10-184-634-48	Sequence 48, App
C 40	37.8	3.0	457	14	US-10-063-685-12	Sequence 12, App
C 41	37.4	2.9	557	14	US-10-184-644-326	Sequence 326, App
C 42	37.4	2.9	557	14	US-10-184-634-326	Sequence 326, App
C 43	37	2.9	186	14	US-10-156-761-5719	Sequence 5719, App
C 44	37	2.9	9025608	14	US-10-156-761-1	Sequence 1, App
C 45	36.8	2.9	1230	12	US-10-882-122A-30075	Sequence 30075, A

```

RESULT 1
US-10-608-504-1
/ Sequence 1, Application US/10608504
/ Publication No. US20040014123A1
/ GENERAL INFORMATION:
/ APPLICANT: KENNERKNECHT, NICOLE
/ APPLICANT: SAHM, HERMANN
/ APPLICANT: EGGELE, LOTMAR
/ APPLICANT: PEFFERLE, WALTER
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
/ TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
/ TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
/ FILE REFERENCE: 7601/8055
/ CURRENT APPLICATION NUMBER: US/10/608,504
/ CURRENT FILING DATE: 2003-06-30
/ PRIOR APPLICATION NUMBER: US 09/471,803
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: DE 199 51 708.8
/ PRIOR FILING DATE: 1999-10-27
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1271
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (101)..(853)
/ OTHER INFORMATION: brnF
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (853)..(1176)
/ OTHER INFORMATION: brnE
/ FEATURE:
/ OTHER INFORMATION: ATCC14752
US-10-608-504-1
Query Match 100.0%; Score 1271; DB 15; Length 1271

```


Tue Apr 20 06:47:23 2004

us-10-608-504-6.rni

Page 8

```

; TELEPHX: (206) 662-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; type: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-03-030-607-12

```

Query Match	2.9%	Score 36.4	DB 3	Length 751
Best Local Similarity	59.8%	Pred. No. 0.099		
Matches 58; Conservative	0	Mismatches 39	Indels 0	Gaps 0

[illegible]

Search completed: April 19, 2004, 18:38:21
Job time : 113.018 secs

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 3205
LENGTH: 1140
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3205

Query Match 2.9%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.097; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGGCCTGATTTCGGCGGCTCCACCGAATGCTGTCATCGCCCTGTTGTGGCGCA 370
DB 465 TACTTCCTGATCAACCTCGCGCGCCCTCGCGCGCTGATCGGCTGACTCTCGCGCTC 406
QY 371 GCGGCCCTGGCGCGCATCGCGCTCACCATTGCTGTGAATTCGCGCAG 422
DB 405 GCGGCCCGCGCGCACCTTCTTGCTACCGCGCTGTCTACTTCTGCTACG 354

RESULT 13
US-09-252-991A-2761
Sequence 2761, Application US/09252991A
Patent No. 6551755
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 2.9%; Score 36.8; DB 4; Length 2061;
Best Local Similarity 58.0%; Pred. No. 0.15; Indels 47; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGGCCTGATTTCGGCGGCTCCACCGAATGCTGTCATCGCCCTGTTGTGGCGCA 370
DB 1273 TACTTCCTGATCAACCTCGCGCGCCCTCGCGCGCTGATCGGCTGACTCTCGCGCTC 1332
QY 371 GCGGCCCTGGCGCGCATCGCGCTCACCATTGCTGTGAATTCGCGCAG 422
DB 1333 GCGGCCCGCGCGCACCTTCTTGCTACCGCGCTGTCTACTTCTGCTACG 1384

RESULT 14
US-09-020-956-12/C
Sequence 12, Application US/09020956
Patent No. 6261562
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-020-956-12

Query Match 2.9%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.099; Indels 39; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 327 CGGGCTCCACCGAATGCTGTCATCGCCCTGTTGTGGCGGAGCGCCCTGGGCGCA 386
DB 410 CGTTCCTTCATCANTCCTCCTCATCTGCTGCTGAGTGCGAGCTGTGTGCGCT 351
QY 387 TCGGGCTCACCACATTGCTGTGAATTCGCGCAGT 423
DB 350 TGTGTACACCAATGCTGAGACCTTCCCTACCT 314

RESULT 15
US-09-030-607-12/C
Sequence 12, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

Db 593 ACGGATGAGCTTGGCGCG 612

RESULT 9

US-09-252-991A-11418

Sequence 11418, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11418

LENGTH: 1086

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11418

Query Match

Best Local Similarity 3.1%; Score 39.2; DB 4; Length 1086;

Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Db 309 TTTCGCGCTGATTTTCGCGGCTCCACCAATGCTGCTATCGCCTCGTTGGCGG 368

222 TGTCCGCACTGCTTCGCGCGCGCCAGTTGCTGCGATCGGCATGCTCAAGGCGG 281

369 CACGCGCCCTGGCGGCGCATGCGGCTCAGCAGATGCTGCTGCTGCTGCTGCTGCT 428

282 GTCCCGGCTTTTGTGATGATCTTCAACACCTCTCTGACCTGCGACGACCTGCTCT 341

429 ATGCGTTTCATTCGCGCTG 448

342 ACGGATGAGCTTGGCGCG 361

Db

RESULT 10

US-09-621-976-8976/c

Sequence 8976, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 8976

LENGTH: 399

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-8976

Query Match

Best Local Similarity 3.1%; Score 39; DB 4; Length 399;

Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

Db 609 GAGTGGGATGCGAGTGTATCTTTGAAATTAAGGCGCTCGAGTCCGCTTTGCT 668

242 SAMRRKKMTCKWGRSSWGRSTGYAMMYKSWCTSRKMMYYKRRKKWRKCTSTKRT 183

669 CTCCTTTGACGCTGACTTGTGATTCCTGCGCAAGAAAAGCAGATCCCTTCTCTGC 728

Db 182 CYRGSTYKCMKAYTKRKRTMTYTYKSYMSYKKTWBMKATAYTKRMKRTKMTW 123

729 TGCTGCGAGTTTGAAGCTTCAACATGCTCTTGATTAATTCAGAGTCCGCTATTGG 788

122 CTWCKMTCTTMMAGTMMYTRRYAKRASKCTSTTTCYCKYAKCKMSYMSMS 63

789 CGCGCTGCTGATCTTCTTGAGTCTGTTGACCATCCGCTACTTCTTGGGAAAAGCT 847

62 MKMKGSMMWKMTYTYTYMKMSKMTYMSMSYARCKWRTYATYTWTCWTGKRT 4

Db

RESULT 11

US-09-621-976-18033

Sequence 18033, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 18033

LENGTH: 474

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 16

OTHER INFORMATION: n=a, g, c or t

US-09-621-976-18033

Query Match

Best Local Similarity 2.9%; Score 37; DB 4; Length 474;

Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

Db 522 CGGCGAGCGCGGAGCTGCTGCGGCTGCGGAGCTTATCTCAAGCAATAGCGTTCACT 581

13 YSGNCCYCCMARKRYSGRMYKSSCYKSCMCMKSYKSGSYKTTTWTW 72

582 CCTACTGAGTATTCGCGGCTCTCACGAGTGGGATGCGAGTGTATTCCTTTGAAA 641

73 WTTTTCGKARRRMSGSKTTTTCGSKKTKSCMAGRWGKYYSRWYCYKAGCYM 132

642 TTAAGGCTGAGTGGCTTGTCTCTCTTTGTCAGCTGAGTGTGATTCCTGCGC 701

133 WKRYCSCCMYTKGGGSMWTTTMMRRRKYKRYTKKKKRTTMMMAACYTTWRS 192

702 GAAGGAAAAGCAATCCCTCTGCTGCTGCTGCGAGTGTGAGCTTCAACATGCTCTTG 761

193 YMMRRRAAAKYYTCMSKTMCAACCCMCMRRASCCMRSCMRBYTMMCYYY 252

762 TGTAAATTCAGGTCAGGCTTATTTGGGCGCTGCTGATCTTTGGGCTGTGACCA 821

253 MMYKGRMYTMMRGMMYRMYYMKKSMWKS CMWKRAWARTKTYTYTAMVYTYKR 312

822 TCCGCTACTCTCTCTGCGGAAAGGCTGTAATAGCAACT 861

313 MCCYMRKTTYCMWMTSRMWSMTARAGAMWMCYTY 352

Db

RESULT 12

US-09-252-991A-3205/c

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

Tue Apr 20 06:47:23 2004

us-10-608-504-6.inl

Page 5

IMMEDIATE SOURCE:
CLONE: PTZEP-F1s
US-08-232-463-14

Query Match 4.1%; Score 52.4; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 2e-06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

560 TCAATGCAAAATGAGCTTCACTGAGGATTCGAGGCTCAGCCGAGTGGCATC 619
1050 TCGAGGAGCTTCCGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1109
620 GCAAGTGTATTCCTTTGAAATTAAGGCTCAGGCTGAGCTTCTTCTTCTTCTC 679
1110 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1169
680 AGCTGACTTGGATTCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 739
1170 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1229
740 TTGAGCTTCACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 799
1230 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1289
800 ATCTTCTTGGGCTTGTGACATCCGATCTTCTTCTTCTTCTTCTTCTTCTTCTT 859
1290 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1349
860 CTGATTTCTCTGATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 919
1350 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1409
920 GGGCGGCTCGCTTCTTAACTTAAAGCCCTTA 951
1410 TTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1441

RESULT 7
US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 893-5007

TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: circular
TOPOLOGY: circular
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-09-107-532A-1266

Query Match 3.1%; Score 39.8; DB 4; Length 768;
Best Local Similarity 49.3%; Pred. No. 0.0072;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

207 TAAAACTTCCCTGCTGCAAGTTTGGGCAATGATCCGATGTTGCTTCTCT 266
92 TCAAGACAGTTTACCTACCGTTTTCGTTATATCGTATTTGACCTTGATTTGATCG 151
267 TGGTTATTCATTAAGGCTAGCAATGGTGGGAGGCCCACTGTTTCCGGCTGATTTTCG 326
152 TTGGAAAGCTGCGGAGTTTATTCATTCACACTAGTGTCAAGTTGATGCTGCTGATG 211
327 CGGGCTCACCGAAATGCTGATCATGCGCTGTTGTGGGCGAGCGCCCTTGGGCGCCA 386
212 CTGGTTCTGCCCATTTATACAGTCAAGTGTGCTGTAACAGCCCATGCTTTTCCA 271
387 TCGGCTCACCATGCTGCTGTAAGTTTCG 417
272 TGGTTTCTGACCTTTCTGATCAATTCGG 302

RESULT 8
US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11376

LENGTH: 1038

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Query Match 3.1%; Score 39.2; DB 4; Length 1038;
Best Local Similarity 55.0%; Pred. No. 0.014;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

309 TTTCCGCTGATTTTTCGGGCTTCACCGAAATGCTGCTATGCTTCTGTTGGGCG 368
473 TGTCCGCTGATGCTTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532
369 CAGCGCCCTGCGGCG 428
533 GTGCGGCTTCTTTTCAATCATCTTCAACACCTTCTCTCACTCGAGACCTGCTCT 592
429 ATGCGTTTCAATTCGCGCTG 448


```

Db 481 TTGGTTTGGCGCTTATCGACGAAGCTACGAGTCACTGCGGCGAGCCGCGAGGCTG 540
Qy 541 GTGGCGGTGGGACTTATCTCAATGCAATAGGCTTTCACCTCCACTGGGTATTCGGCGG 600
Db 541 GTGGCGGTGGGACTTATCTCAATGCAATAGGCTTTCACCTCCACTGGGTATTCGGCGG 600
Qy 601 TCTGACCGGAGTGGCGATCGCAGAGTTGATCTCTTTTGAATTAAGGAGCTCGAGTTGCG 660
Db 601 TCTGACCGGAGTGGCGATCGCAGAGTTGATCTCTTTTGAATTAAGGAGCTCGAGTTGCG 660
Qy 661 CTTTGTCTCTCTTTGTGACGCTGACTTGAATTCCTGCCGAAAGAAAAGCATGCC 720
Db 661 CTTTGTCTCTCTTTGTGACGCTGACTTGAATTCCTGCCGAAAGAAAAGCATGCC 720
Qy 721 TTCTGTGCTGCTGCGAGGTTGAGGTTGACCATGCTCTTGAGTAATTCAGAGTCAAGC 780
Db 721 TTCTGTGCTGCTGCGAGGTTGAGGTTGACCATGCTCTTGAGTAATTCAGAGTCAAGC 780
Qy 781 CCTATTGCGGCGCTGCTGATCTTCTTGAGTCTGTTGACATCCGATCTTCTTGGG 840
Db 781 CCTATTGCGGCGCTGCTGATCTTCTTGAGTCTGTTGACATCCGATCTTCTTGGG 840
Qy 841 AAAGCTGCTAAATGACAACTGATTTCTCCCTGATTCCTCCCTGTTGTCGAGTATGCA 900
Db 841 AAAGCTGCTAAATGACAACTGATTTCTCCCTGATTCCTCCCTGTTGTCGAGTATGCA 900
Qy 901 GTCAATTACTTTTGCGCTCGGCGGCTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
Db 901 GTCAATTACTTTTGCGCTCGGCGGCTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
Qy 961 CAATTGTTGGGCAAAATGGCGATGTCGATGCGCAGAGAAATCCTTGCCATTTTACCGCA 1020
Db 961 CAATTGTTGGGCAAAATGGCGATGTCGATGCGCAGAGAAATCCTTGCCATTTTACCGCA 1020
Qy 1021 TCAAGCTTTCGAGCAATGCGATAGATCTGAAGACTTAACCTTGGTCTGATGCCGT 1080
Db 1021 TCAAGCTTTCGAGCAATGCGATAGATCTGAAGACTTAACCTTGGTCTGATGCCGT 1080
Qy 1081 GCGATTACAGTGTGTGGCGCATCTTCTTGCGGCTGACGCACTTGTGAGCGTTGGCGCT 1140
Db 1081 GCGATTACAGTGTGTGGCGCATCTTCTTGCGGCTGACGCACTTGTGAGCGTTGGCGCT 1140
Qy 1141 GGCACCATGCTTTTGTGTGAAGTCTGTAATCTTAACTGATTAATTAAGAAAT 1200
Db 1141 GGCACCATGCTTTTGTGTGAAGTCTGTAATCTTAACTGATTAATTAAGAAAT 1200
Qy 1201 CCGCATGCGCTCAATTGTAAGGGGATGCGGATTTTGAAGAACTGAAAAGGCTTAAG 1260
Db 1201 CCGCATGCGCTCAATTGTAAGGGGATGCGGATTTTGAAGAACTGAAAAGGCTTAAG 1260
Qy 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

```

```

RESULT 3
US-09-471-803A-2
; Sequence 2, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMAN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS. PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/263496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27

```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1.
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: brnF
; OTHER INFORMATION: ATCC14752
US-09-471-803A-2

Query Match      59.2%; Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred.No. 6.6e-243;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 GTGCAAAAACGCAAGATTTATTCAGACCTGAGAGTGTGGCATTCAGAGGAGAGCCCTG 160
Db 1 GTGCAAAAACGCAAGATTTATTCAGACCTGAGAGTGTGGCATTCAGAGGAGAGCCCTG 60
Qy 161 GAACGAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAAGCTCCCT 220
Db 61 GAACGAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAAGCTCCCT 120
Qy 221 GCTGCAAGTTTGGGCAATGACCCGATGATGATGATGATGATGATGATGATGATGATG 280
Db 121 GCTGCAAGTTTGGGCAATGACCCGATGATGATGATGATGATGATGATGATGATGATG 180
Qy 281 GGCTAGAAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
Db 181 GGCTAGAAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 341 ATGCTGATATGAGCTGCTGTTGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
Db 241 ATGCTGATATGAGCTGCTGTTGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 401 TTGCTGATGAACTTCGCGAGAGATTTCTATGAGGTTTCAATCCGCTGATGATGATGAA 460
Db 301 TTGCTGATGAACTTCGCGAGAGATTTCTATGAGGTTTCAATCCGCTGATGATGATGAA 360
Qy 461 AACCCATGCGCGGTTCTATTCGTTTGGCGCTTATGAGAGAGAGAGAGAGAGAGAGAG 520
Db 361 AACCCATGCGCGGTTCTATTCGTTTGGCGCTTATGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 521 GCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 580
Db 421 GCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 581 TCTTACTGAGTATTCGCGCGCTCTACCGAGAGTGCAGATGCAAGATGATTCCTTTGAA 640
Db 481 TCTTACTGAGTATTCGCGCGCTCTACCGAGAGTGCAGATGCAAGATGATTCCTTTGAA 540
Qy 641 ATTAAGGCGCTGAGATTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
Db 541 ATTAAGGCGCTGAGATTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Qy 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATGCTCTT 760
Db 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATGCTCTT 660
Qy 761 GTGTAATTCAGAGTCAAGGCGCTATTTGCGGCGCTGCTGATCTTCTTGAGTCTGTTGAC 820
Db 661 GTGTAATTCAGAGTCAAGGCGCTATTTGCGGCGCTGCTGATCTTCTTGAGTCTGTTGAC 720
Qy 821 ATCCGATCTTCTTGGGAAAGGCTGCTAA 853
Db 721 ATCCGATCTTCTTGGGAAAGGCTGCTAA 753

```

```

RESULT 4
US-09-471-803A-4
; Sequence 4, Application US/09471803A

```

```

Db      181  TCGGGGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCTTGTGAGGTTGGGCACTGA 240
      241  CCCGATGGATTTGGGTTGGTCTCTGGTTATTCATATACGGCTACGATGGTGGGACAGC 300
      241  CCCGATGGATTTGGGTTGGTCTCTGGTTATTCATATACGGCTACGATGGTGGGACAGC 300
      301  CCCACTGTTTTCCGGCTGATATTTTGGCGGGCTCCACCGAATGCTGTCATGCGCTTCTGT 360
      301  CCCACTGTTTTCCGGCTGATATTTTGGCGGGCTCCACCGAATGCTGTCATGCGCTTCTGT 360
      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGCGCTACACCATGCTGTGAATTCCTCGGCA 420
      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGCGCTACACCATGCTGTGTGAATTCCTCGGCA 420
      421  CGTATTCATAGCGTTTCAATTCGCGTGCATGTGGTCAAAAACCCCATTTGCCGTTCTTA 480
      421  CGTATTCATAGCGTTTCAATTCGCGTGCATGTGGTCAAAAACCCCATTTGCCGTTCTTA 480
      481  TTGCGTTTCCGGCTTATGAGAGAAAGCTACAGCATGCTGGGCGAGGCGCGAGGCTG 540
      481  TTGCGTTTCCGGCTTATGAGAGAAAGCTACAGCATGCTGGGCGAGGCGCGAGGCTG 540
      481  TTGCGTTTCCGGCTTATGAGAGAAAGCTACAGCATGCTGGGCGAGGCGCGAGGCTG 540
      541  GTTCGGGCTGCGCATTTATCTCAATGCAATAGCGTTTCACTGCGGATTTGGCGCG 600
      541  GTTCGGGCTGCGCATTTATCTCAATGCAATAGCGTTTCACTGCGGATTTGGCGCG 600
      601  TCTCACCGGAGTGGCGCATGGCAGAGTGAATTCCTTTGAAATTAAGGCGCTGAGATGGC 660
      601  TCTCACCGGAGTGGCGCATGGCAGAGTGAATTCCTTTGAAATTAAGGCGCTGAGATGGC 660
      661  CCTTCTCTCTCTTGTGACGCTGACTTGTGAATTCCTGCGCAGAAAGACAGATCCC 720
      661  CCTTCTCTCTCTTGTGACGCTGACTTGTGAATTCCTGCGCAGAAAGACAGATCCC 720
      721  TTCTCTGCTGCTGGCGAGTTTGAAGCTTCAACATTCCTTGTGTGATTCGAGTCAAGC 780
      721  TTCTCTGCTGCTGGCGAGTTTGAAGCTTCAACATTCCTTGTGTGATTCGAGTCAAGC 780
      781  CCTATTGCGGCGCTGCTGATCTCTTGGGTCTGTGACATCCGCTACTTCTTCTTGGG 840
      781  CCTATTGCGGCGCTGCTGATCTCTTGGGTCTGTGACATCCGCTACTTCTTCTTGGG 840
      841  AAAGGCTGCTAAATGACACTGATTTCTCTGTATTCCTTGTGTGACATGATGCA 900
      841  AAAGGCTGCTAAATGACACTGATTTCTCTGTATTCCTTGTGTGACATGATGCA 900
      901  GTGATTAATTTTGGGCTCGGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTAAGTGAATCA 960
      901  GTGATTAATTTTGGGCTCGGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTAAGTGAATCA 960
      901  GTGATTAATTTTGGGCTCGGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTAAGTGAATCA 960
      961  CAATTTGTGGGCAAAATGCGATGTGAGTGCAGACAGAAATCTTGGCAATTTTACCGCA 1020
      961  CAATTTGTGGGCAAAATGCGATGTGAGTGCAGACAGAAATCTTGGCAATTTTACCGCA 1020
      1021  TCAACGTTTGGCGCAATGCGATGATCTGAAGACCTTAACCTTGTGCTCAATGGCGCT 1080
      1021  TCAACGTTTGGCGCAATGCGATGATCTGAAGACCTTAACCTTGTGCTCAATGGCGCT 1080
      1081  GCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTGGCGCT 1140
      1081  GCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTGGCGCT 1140
      1141  GCGACCAATGCTTTTGTGAGACTGTGAATCTTTTCTTAATCTGCAATTAATCAAAAT 1200
      1141  GCGACCAATGCTTTTGTGAGACTGTGAATCTTTTCTTAATCTGCAATTAATCAAAAT 1200
      1201  CCGCATGCGCTCAATTTGAAGGGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
      1201  CCGCATGCGCTCAATTTGAAGGGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
      1261  CAGACAGCGCT 1271
  
```

```

Db      1261  CAGACAGCGCT 1271

RESULT 2
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNEDY, NICOLE
; APPLICANT: SAHM, HERMAN
; APPLICANT: EGGELING, LOTAR
; APPLICANT: PEEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MMS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: DIMP
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: DIMP
; OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match      100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1  GCGCATCAATGAGATCTAGCTTCAATATTTGACATATAGCTTGAAGTGGGCAAC 60
      1  GCGCATCAATGAGATCTAGCTTCAATATTTGACATATAGCTTGAAGTGGGCAAC 60
      61  TGGCAACAAATCTACCGGCAATTTGTGATGATTTGATGATGATGATGATGATGAT 120
      61  TGGCAACAAATCTACCGGCAATTTGTGATGATTTGATGATGATGATGATGATGAT 120
      121  TCATTCAAGCTTGAAGGTGTGCGCATCAAGGAGCGCTGGAACCAAGATGATAAGTTA 180
      121  TCATTCAAGCTTGAAGGTGTGCGCATCAAGGAGCGCTGGAACCAAGATGATAAGTTA 180
      181  TGGGCGCTACGAAATTCGCGCAAGGTCTPAAAAACCTCCCTTGTGACAGTTTGGCATGA 240
      181  TGGGCGCTACGAAATTCGCGCAAGGTCTPAAAAACCTCCCTTGTGACAGTTTGGCATGA 240
      241  CCCGATGATTTGCGGCTGATTTTGGCGGGCTCCACCGAATGCTGATCGCCCTGCT 300
      241  CCCGATGATTTGCGGCTGATTTTGGCGGGCTCCACCGAATGCTGATCGCCCTGCT 300
      301  CCCACTGTTTTCCGGCTGATATTTTGGCGGGCTCCACCGAATGCTGATCGCCCTGCT 360
      301  CCCACTGTTTTCCGGCTGATATTTTGGCGGGCTCCACCGAATGCTGATCGCCCTGCT 360
      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGCGCTACACCATGCTGTGAATTCCTCGGCA 420
      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGCGCTACACCATGCTGTGTGAATTCCTCGGCA 420
      421  CGTATTCATAGCGTTTCAATTCGCGTGCATGTGGTCAAAAACCCCATTTGCCGTTCTTA 480
      421  CGTATTCATAGCGTTTCAATTCGCGTGCATGTGGTCAAAAACCCCATTTGCCGTTCTTA 480
      481  TTGCGTTTCCGGCTTATGAGAGAAAGCTACAGCATGCTGGGCGAGGCGCGAGGCTG 540
      481  TTGCGTTTCCGGCTTATGAGAGAAAGCTACAGCATGCTGGGCGAGGCGCGAGGCTG 540
  
```


Tue Apr 20 06:47:23 2004

us-10-608-504-6.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 19, 2004, 10:57:33 ; Search time 107.018 seconds

(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-6

Perfect score: 1271

Sequence: 1 gcgcgacatcagatcagcgtg.....agccttaagcagacgcgcct 1271

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
- 2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
- 3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
- 4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
- 5: /cgn2_6/prodata/2/ina/PCTUS.COMB.seq:*
- 6: /cgn2_6/prodata/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1271	100.0	1271	US-09-471-803A-1	Sequence 1, Appli
2	1271	100.0	1271	US-09-471-803A-6	Sequence 6, Appli
3	753	59.2	753	US-09-471-803A-2	Sequence 2, Appli
4	324	25.5	324	US-09-471-803A-4	Sequence 4, Appli
5	58.6	4.6	756	US-09-489-039A-3016	Sequence 3016, Ap
6	52.4	4.1	7218	US-08-232-463-14	Sequence 14, Appli
7	39.8	3.1	768	US-09-107-532A-1266	Sequence 1266, Ap
8	39.2	3.1	1038	US-09-252-991A-11376	Sequence 11376, A
9	39.2	3.1	1086	US-09-252-991A-11418	Sequence 11418, A
10	39.2	3.1	399	US-09-621-976-8976	Sequence 8976, Ap
11	37	2.9	474	US-09-621-976-18033	Sequence 18033, A
12	36.8	2.9	1140	US-09-252-991A-3205	Sequence 3205, Ap
13	36.8	2.9	2061	US-09-252-991A-2761	Sequence 2761, Ap
14	36.4	2.9	751	US-09-020-956-12	Sequence 12, Appli
15	36.4	2.9	751	US-09-030-907-12	Sequence 12, Appli
16	36.4	2.9	751	US-09-439-313-12	Sequence 12, Appli
17	36.4	2.9	751	US-09-352-616A-12	Sequence 12, Appli
18	36.4	2.9	751	US-09-352-616A-12	Sequence 12, Appli
19	36.4	2.9	751	US-09-159-812-12	Sequence 12, Appli
20	36.4	2.9	751	US-09-636-215-12	Sequence 12, Appli
21	36.4	2.9	751	US-09-685-166A-12	Sequence 12, Appli
22	36.4	2.9	751	US-09-115-453-12	Sequence 12, Appli
23	36.4	2.9	751	US-09-688-489-12	Sequence 12, Appli
24	35.4	2.8	430	US-09-621-976-16656	Sequence 16656, A
25	35.2	2.8	801	US-09-328-352-2399	Sequence 2399, Ap
26	35.2	2.8	4403765	US-09-103-840A-2	Sequence 2, Appli
27	35.2	2.8	4411529	US-09-103-840A-1	Sequence 1, Appli

28	34.8	2.7	867	US-09-489-039A-5047	Sequence 5047, Ap
29	34.6	2.7	666	US-08-875-034A-1	Sequence 1, Appli
30	34.4	2.7	744	US-09-489-039A-1307	Sequence 1307, Ap
31	34.4	2.7	1224	US-09-266-965-22	Sequence 22, Appli
32	34.4	2.7	12249	US-09-266-965-74	Sequence 74, Appli
33	34.4	2.7	18331	US-09-266-965-96	Sequence 96, Appli
34	34.2	2.7	594	US-09-252-991A-13729	Sequence 13729, A
35	34.2	2.7	678	US-09-252-991A-13832	Sequence 13832, A
36	33.4	2.6	1875	US-08-422-869-21	Sequence 21, Appli
37	33.4	2.6	2949	US-08-433-522A-1	Sequence 1, Appli
38	33.4	2.6	2949	US-09-135-166-1	Sequence 1, Appli
39	33.4	2.6	2949	US-08-942-046-1	Sequence 1, Appli
40	33.4	2.6	2950	US-08-433-522A-5	Sequence 5, Appli
41	33.4	2.6	2950	US-09-135-166-5	Sequence 5, Appli
42	33.4	2.6	2984	US-08-942-046-5	Sequence 5, Appli
43	33.4	2.6	2984	US-08-433-522A-3	Sequence 3, Appli
44	33.4	2.6	2984	US-09-135-166-3	Sequence 3, Appli
45	33.4	2.6	2984	US-08-942-046-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1					
US-09-471-803A-1					
Sequence 1, Application US/09471803A					
Patent No. 6613545					
GENERAL INFORMATION:					
APPLICANT: KENNERNECHT, NICOLE					
APPLICANT: SAM, HERMANN					
APPLICANT: EGGELING, LOTHAR					
APPLICANT: PFEFFERLE, WALTER					
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF					
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE					
FILE REFERENCE: 21123/265496/MAS					
CURRENT APPLICATION NUMBER: US/09/471,803A					
CURRENT FILING DATE: 1999-12-23					
PRIOR APPLICATION NUMBER: DE 199 51 708.8					
PRIOR FILING DATE: 1999-10-27					
NUMBER OF SEQ ID NOS: 12					
SOFTWARE: Patentm Ver. 2.1					
SEQ ID NO 1					
LENGTH: 1271					
TYPE: DNA					
ORGANISM: Corynebacterium glutamicum					
FEATURE:					
NAME/KEY: gene					
LOCATION: (101)..(853)					
OTHER INFORMATION: brnF					
NAME/KEY: gene					
LOCATION: (853)..(1176)					
OTHER INFORMATION: brnE					
OTHER INFORMATION: ATCC14752					
US-09-471-803A-1					
Query Match					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	GC	GC	GC	GC
DB	1	GC	GC	GC	GC
QY	61	TG	GC	GC	GC
DB	61	TG	GC	GC	GC
QY	121	TC	AT	CA	GC
DB	121	TC	AT	CA	GC
QY	181	TC	GC	GC	GC
DB	181	TC	GC	GC	GC

```
Db 233 TCAATGTAATTTCCGACATTTTCTACGCTCACCCTTCCCGCCACCGCATCAAGT 232
QY 462 ACCCATTTGCCGTTTCTATTGGGTTTGGCCTTATGACGAAGCCTAGCACTCACTG 521
Db 293 CCGGCGCGCGCCGCGCCTATTCCACCTACGCGCTTACGACAGCTCCTACGCCATCTGT 352
QY 522 CGGCGAGGCCGCGAGGCTGTGGGCTG---GCCACTATCTCAATGGAATAGCCTTTC 578
Db 353 CAGCCCGCCACCTGCGATATCACTGACGCGGCTCTTACCGTTCAATTTTGTGCC 412
QY 579 ACTCTACTGGGATTTCCGCGGCTCACCGGAGTGGCGATGCAAGTTGATTCCTTTG 638
Db 413 AAGCTGTGGGTATCCAGGAATATTGGCGCCTTGTGTGTCGAAGTGTGCCGATG 472
QY 639 AAATTAAAGGCTCGAGTTGCCCTTGTCTCTTTGTCAAGCTGACTTTGATTCCT 698
Db 473 ATCTAAAGGATGATTTTGCCCTGACCGCGCTGTGTGTGCTGCGTGGAGGCAAT 532
QY 699 GCCGAACGAAAGCAGATCCCTTCTGTGCTGCTGCAAGTTTGAGCTTCAACATTTGCTC 758
Db 533 TCAAAATTAACAGATTAATTCGCGCATTAATGCGGTGATATGGCTCTGTTCCG 592
QY 759 TTGTGTAAATTCAGGTCAGGCCCTAATTTGGGCGCTGTGATCTTCTTGGGTCTGTGA 818
Db 593 GTTTGTGGCGCCCGAGCAGATGCTGTATCGCTTTGACCACGATCTTTTGAATCCTTC 652
QY 819 CCATCCGGTACTTCTTC 835
Db 653 TTCTCCGCGTCCGCTTC 669
```

Search completed: April 19, 2004, 11:42:05
Job time : 593.9 secs

XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KJOW) KYOMA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI: 2001-376931/40.
 DR P-PSDB: AAG93154.
 XX
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 3408; 246bp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences, from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
 CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 10.0%; Score 127.4; DB 5; Length 711;
 Best Local Similarity 51.5%; Pred. No. 9e-30;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 222 CTGCAAGTTGGGAGATGACCGATTTGGTATGCTCTGTTATTCATATACG 281
 DB 53 CGGTGGTTGGGAGTATTCGCTGGTGGTGGCTTTGGCTGTATGATGATGATGATGAT 112
 QY 282 GCTACGAATGGTGGGAGCCCACTGTTTCGCGCTGATTTTGGCGGCTCCACCGAA 341
 DB 113 GTTTCGCTGGTGGTGGAGCGCGGATTTCTCTTGATGATGATGATGATGATGAT 172
 QY 342 TGCCTGATATGCGCTGCTGTTGGGCGAGCGCCCTGGGCGCATGCGCTCCACCAT 401
 DB 173 TTTCTGCAATCGGCAATGATACCGAGATGCGCTGCTTTGGCGGCGGCTGCTGTT 232
 QY 402 TGTCTGTAATCTCCGCAAGTATTCATGCTTTTCATCCCGCTGCATGTGTCAAA 461
 DB 223 TCAATGTAATTTCCGCAATTTCTACGATCTCACTTCCACGCAAGCATCAAGT 292
 QY 462 ACCCATATGCGCTTTCTATTTGCTTTGCGCTTATGAGAGGACCAAGCATGCTG 521
 DB 293 CGGCGCGCGCGCGCTTATTCACCTTACGCGCTTACGAGAGTCTTACGCAATGCTGT 352
 QY 522 CGGCAAGGCGCGCAGAGCTGTGCGGCTG--GCGACTTATTCATGCAATGAGCGTTTC 578
 DB 353 CAGCGCGCGCACCTGCGCATATCAGTGGAGCGCGGCTGCTTACGTTCAAAATTTGTGCC 412
 QY 579 ACTCTACCTGGGATTTGCGCGCTTCAACGAGTGGCATGCGCAAGTGAATCTTTTG 638
 DB 413 AAGCTCTGGGGTTATTCAGAGAAATTTGCGCGCTTGTGTCAGTGTGCGCGCATG 472
 QY 639 AAATTAAGGCGCTTCAGTTGCGCTTTGCTCTCTTTGTCACGCTGACTTTGATTCCT 698
 DB 473 ATCTAAAGAGCATGATTTTGGCTGACCGGCTGTTTGTGCTGCGGATGAGGAGCAT 532
 QY 699 GCGGAAGCAAAAGCATGCTTCTCTGCTGCTGCGAGTTCAGCTTACCATTTGCTC 758
 DB 533 TCAAAATTAACAGGATTTATGCTGCGCATTTTGGCGGTGATTTGGCTGCTGCTTCCG 592

QY 759 TTGTGTAATTCAGATGAGCCCTATTTGCGCGCTGCTGATCTTTGGTCTGTGA 818
 DB 593 GTTTTGTGGCGCCGACAGAGATGCTGTATGCTTTGACCAAGTATTTGATCTTC 652
 QY 819 CCATCGGATCTTCTTC 835
 DB 653 TTCTCGGCTCGGCTTC 669
 RESULT 15
 ID ACA01294
 XX ACA01294 standard; DNA; 711 BP.
 AC ACA01294;
 XX
 DT 03-JUN-2003 (first entry)
 DE
 XX C. glutamicum derived ORF SEQ ID 1285.
 DE
 XX Corynebacterium; nucleic acid array; fermentation; culture; ds.
 KM
 XX Corynebacterium glutamicum.
 OS
 XX DE10128510-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 13-JUN-2001; 2001DE-01028510.
 XX
 PR 13-JUN-2001; 2001DE-01028510.
 XX
 PA (DEGS) DEGUSSA AG.
 PI Farwick M, Moessel B, Pfeifferle W, Bathe B, Hutmacher K;
 XX
 DR WPI: 2003-279970/28.
 XX
 PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 XX
 PS Claim 1; Page 456; 709bp; German.
 XX
 CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA0010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 10.0%; Score 127.4; DB 7; Length 711;
 Best Local Similarity 51.5%; Pred. No. 9e-30;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 222 CTGCAAGTTGGGAGATGACCGATTTGGTATGCTCTGTTATTCATATACG 281
 DB 53 CGGTGGTTGGGAGTATTCGCTGGTGGTGGCTTTGGCTGTATGATGATGATGATGAT 112
 QY 282 GCTACGAATGGTGGGAGCCCACTGTTTCGCGCTGATTTTGGCGGCTCCACCGAA 341
 DB 113 GTTTCGCTGGTGGTGGAGCGCGGATTTCTCTTGATGATGATGATGATGATGAT 172
 QY 342 TGCCTGATATGCGCTGCTGTTGGGCGAGCGCCCTGGGCGCATGCGCTCCACCAT 401
 DB 173 TTTCTGCAATCGGCAATGATACCGAGATGCGCTGCTTTGGCGGCGGCTGCTGTT 232
 QY 402 TGTCTGTAATCTCCGCAAGTATTCATGCTTTTCATTTCCGCTGCATGTGTCAAA 461

```

FT      -10_signal      /*tag= a
FT      88..93
FT      /*tag= b
FT      151..615
FT      CDS              /*tag= c
FT      /product= "ltp"
FT
XX      EP1090993-A1.
XX      11-APR-2001.
XX      29-SEP-2000; 2000EP-00121159.
XX      05-OCT-1999; 99DE-01047792.
XX      (DEGS ) DEGUSA-HUELS AG.
XX
PI      Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX      WPI, 2001-292927/31.
XX      P-PSDB; AAB70881.
XX
PT      New ltp gene from coryneform bacteria, used to prepare transformants with
PT      increased synthesis of amino acids, particularly lysine and isoleucine.
XX
PS      Claim 4; Page 14-15; 22pp; German.
XX
CC      This invention describes a novel isolated nucleic acid (I) from
CC      coryneform bacteria used for the fermentative production of selected L-
CC      amino acids, by fermenting the amino acid-producing coryneform in which
CC      at least the ltp gene has been weakened or amplified, then isolating
CC      amino acids that have accumulated in the medium or cells. (I) is used to
CC      transform coryneforms for production of L-amino acids, specifically
CC      lysine and isoleucine, which are used in medicine and particularly as
CC      animal feed supplement. It may also be used as probes and primers for
CC      isolating related sequences. Regulating expression of (I) improves
CC      production of amino acids, especially of L-lysine. This sequence encodes
CC      the Corynebacterium glutamicum ltp protein which is used in the method
CC      described in the invention
XX
SQ      Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;
XX
Query Match      14.6%; Score 185; DB 4; Length 715;
Best Local Similarity 100.0%; Pred. No. 2.9e-48;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      1 GCGGATCATGGAATCTAGCTCATATATTGCAATAGCTAGTTGAGTCCGCAAC 60
DB      185 GCGGATCATGGAATCTAGCTCATATATTGCAATAGCTAGTTGAGTCCGCAAC 126
QY      61 TGGCAACAAAATCTACCGGCAATTTGTGATGATTTGATGTCGCAAAAAACGCAAGAT 120
DB      125 TGGCAACAAAATCTACCGGCAATTTGTGATGATTTGATGTCGCAAAAAACGCAAGAT 66
QY      121 TCATTCAAGCTTGAGGTGTGCGATCCAAAGGAGCCCTGGAACGAGATTAAGGTTA 180
DB      65 TCATTCAAGCTTGAGGTGTGCGATCCAAAGGAGCCCTGGAACGAGATTAAGGTTA 6
QY      181 TCGGC 185
DB      5 TCGGC 1
XX
RESULT 13
ACA01968
ID      ACA01968 standard; DNA; 177 BP.
XX
AC      ACA01968;
XX
DT      04-JUN-2003 (first entry)
XX
DE      C. glutamicum derived ORF SEQ ID 1959.
XX

```

```

KM      Coryneform; nucleic acid array; fermentation; culture; ds.
XX
XX      Corynebacterium glutamicum.
XX      OS
XX      DE10128510-A1.
XX      PN
XX      19-DEC-2002.
XX      PD
XX      13-JUN-2001; 2001DE-01028510.
XX      PF
XX      13-JUN-2001; 2001DE-01028510.
XX      PR
XX      13-JUN-2001; 2001DE-01028510.
XX      PA
XX      (DEGS ) DEGUSA AG.
XX      PI
XX      Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;
XX      WPI; 2003-279970/28.
XX
PT      New nucleic acid array useful for monitoring mRNA expression of
PT      Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT      from Corynebacterium glutamicum.
XX
PS      Claim 1; Page 647; 709pp; German.
XX
CC      This invention describes a novel nucleic acid array involving
CC      Corynebacterium glutamicum polynucleotides. The arrays are used to
CC      analyse C. glutamicum, particularly for monitoring a fermentation process
CC      to determine expression levels of C. glutamicum cellular mRNA. Such
CC      monitoring particularly differentiates between expression levels of
CC      different strains of C. glutamicum and allows the adjustment of different
CC      culture and fermentation conditions. ACA0010-ACA02188 represent C.
CC      glutamicum derived polynucleotides described in the disclosure of the
CC      invention
XX
SQ      Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;
XX
Query Match      13.9%; Score 177; DB 7; Length 177;
Best Local Similarity 100.0%; Pred. NO. 4.8e-46;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      30 TTGCATATAGCTAGTTGAGTGGCAACTGGCAACAAATACCCGGCAATTTGTG 89
DB      1 TTGCATATAGCTAGTTGAGTGGCAACTGGCAACAAATACCCGGCAATTTGTG 60
QY      90 ATGATTTAGTGTGCAAAAAACGCAAGAGATTCTTCAAGCTTGAGTGTGCGCATCCA 149
DB      61 ATGATTTAGTGTGCAAAAAACGCAAGAGATTCTTCAAGCTTGAGTGTGCGCATCCA 120
QY      150 AGGAGCCCTGGAACCAAGATTAAGGTATTCGCGCTACGAATCGCGCAAGTTC 206
DB      121 AGGAGCCCTGGAACCAAGATTAAGGTATTCGCGCTACGAATCGCGCAAGTTC 177
XX
RESULT 14
AAH68373
ID      AAH68373 standard; DNA; 711 BP.
XX
AC      AAH68373;
XX
DT      26-SEP-2001 (first entry)
XX
DE      C glutamicum coding sequence fragment SEQ ID NO: 3408.
XX
XX      Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX      organic acid synthesis; ds.
XX      KM
XX      Corynebacterium glutamicum.
XX      OS
XX      EP1108790-A2.
XX      PN
XX      20-JUN-2001.
XX      PD
XX      18-DEC-2000; 2000EP-00127688.
XX

```

XX WPI: 2001-376931/40.
 DR P-PSDB: AAG90036.
 PT Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 290; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* bacterium *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC *Corynebacterium* bacterium, and identifying a homologue of a gene derived from
 CC *Corynebacterium* bacterium. *Corynebacterium* bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
 XX
 Query Match 25.5%; Score 324; DB 5; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.4e-93;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 853 ATGACACTGATTTCTCCGTATTCCTGTTGTCGAGTATGCGACTTACTTT 912
 DB 1 ATGACACTGATTTCTCCGTATTCCTGTTGTCGAGTATGCGACTTACTTT 60
 QY 913 GCGCTCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGAATCACAATTGCGG 972
 DB 61 GCGCTCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGAATCACAATTGCGG 120
 QY 973 AAAATGGCGATGATGATGCGAGAGATCCTTGCCATTGACCGCATCAAGTTTGGC 1032
 DB 121 AAAATGGCGATGATGATGCGAGAGATCCTTGCCATTGACCGCATCAAGTTTGGC 180
 QY 1033 AGCAATGGATAGATCTGAAGACTTAACCTTGTCTGATGCGGTTSCGATTACAGTG 1092
 DB 181 AGCAATGGATAGATCTGAAGACTTAACCTTGTCTGATGCGGTTSCGATTACAGTG 240
 QY 1093 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTGAGCGTTGGCGCTGGCAACATCGTT 1152
 DB 241 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTGAGCGTTGGCGCTGGCAACATCGTT 300
 QY 1153 TTTGTTGAGCTGTGGAATCTTTTC 1176
 DB 301 TTTGTTGAGCTGTGGAATCTTTTC 324
 XX
 RESULT 11
 ID ACA01970 standard; DNA; 324 BP.
 AC ACA01970;
 AC ACA01970;
 AC 04-JUN-2003 (first entry)
 DT
 XX C. glutamicum derived ORF SEQ ID 1961.
 DE
 XX Coryneform; nucleic acid array; fermentation; culture; ds.
 KM
 OS Corynebacterium glutamicum.
 XX
 PN DE10128510-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 13-JUN-2001; 2001DB-01028510.

XX 13-JUN-2001; 2001DB-01028510.
 PR
 XX (DEGS) DEGUSA AG.
 XX
 PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Huthmacher K;
 XX WPI: 2003-279970/28.
 DR
 XX
 PT New nucleic acid array useful for monitoring mRNA expression of
 PT *Corynebacterium glutamicum* during fermentation, comprising nucleic acid
 PT from *Corynebacterium glutamicum*.
 XX
 PS Claim 1; Page 647; 709pp; German.
 XX
 CC This invention describes a novel nucleic acid array involving
 CC *Corynebacterium glutamicum* polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention
 XX
 SQ Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
 XX
 Query Match 25.5%; Score 324; DB 7; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.4e-93;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 853 ATGACACTGATTTCTCCGTATTCCTGTTGTCGAGTATGCGACTTACTTT 912
 DB 1 ATGACACTGATTTCTCCGTATTCCTGTTGTCGAGTATGCGACTTACTTT 60
 QY 913 GCGCTCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGAATCACAATTGCGG 972
 DB 61 GCGCTCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGAATCACAATTGCGG 120
 QY 973 AAAATGGCGATGATGATGCGAGAGATCCTTGCCATTGACCGCATCAAGTTTGGC 1032
 DB 121 AAAATGGCGATGATGATGCGAGAGATCCTTGCCATTGACCGCATCAAGTTTGGC 180
 QY 1033 AGCAATGGATAGATCTGAAGACTTAACCTTGTCTGATGCGGTTSCGATTACAGTG 1092
 DB 181 AGCAATGGATAGATCTGAAGACTTAACCTTGTCTGATGCGGTTSCGATTACAGTG 240
 QY 1093 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTGAGCGTTGGCGCTGGCAACATCGTT 1152
 DB 241 GTGGCGCATCTTCTTGCGCGGTGACGCACTTGTGAGCGTTGGCGCTGGCAACATCGTT 300
 QY 1153 TTTGTTGAGCTGTGGAATCTTTTC 1176
 DB 301 TTTGTTGAGCTGTGGAATCTTTTC 324
 XX
 RESULT 12
 ID AAF61688/C
 AC AAF61688 standard; DNA; 715 BP.
 AC AAF61688;
 AC 12-JUN-2001 (first entry)
 DT
 XX C. glutamicum lrp encoding DNA.
 DE
 XX Fermentation; L-lysine acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN Key -35_signal Location/Qualifiers
 XX 62.67
 XX
 PF

Db 376 AATGSAATCTAGCTTATATATTGCAATACCTAGTTGAGGTGCGCAACTGGCACA 317

Qy 69 AAACCTACCCGGCAATTGTGTGATGATTTGATGTGCAAAAACGCAAGATTCATTCAA 128

Db 316 AAACCTACCCGGCAATTGTGTGATGATTTGATGTGCAAAAACGCAAGATTCATTCAA 257

Qy 129 GCGTGAAGGTGTGGCATCCAGAGGACCCCTGGAAACGATGATTAAGGTATCGGGCT 188

Db 256 GCGTGAAGGTGTGGCATCCAGAGGACCCCTGGAAACGATGATTAAGGTATCGGGCT 197

Qy 189 ACGAAATCGGCAGAGTCTAAAAACCTCCCTGCTGACAGTTTGAGCAATGACCGAATTG 248

Db 196 ACGAAATCGGCAGAGTCTAAAAACCTCCCTGCTGACAGTTTGAGCAATGACCGAATTG 137

Qy 249 GTAATTGGTTGGTCTCTTGTTATTCATATACGGCTAAGAAAGGTGGGCAAGCCCACTGT 308

Db 136 GTATTGTGGTTGGTCTCTTGTTATTCATATACGGCTAAGAAAGGTGGGCAAGCCCACTGT 77

Qy 309 TTTTCGGGCTGATTTTTCGGGGCTCCACGGAATGCTGATCGCCCTCGTTTGAGGCG 368

Db 76 TTTTCGGGCTGATTTTTCGGGGCTCCACGGAATGCTGATCGCCCTCGTTTGAGGCG 17

Qy 369 CAGGCCCCCTGG 380

Db 16 CAGGCCCCGGG 5

RESULT 9
AAH21111
ID AAH21111 standard; DNA; 324 BP.

DT	05-SEP-2001	(first entry)
XX		
DE	C. glutamicum brne DNA.	

KM L-amino acid production; bnf; bnf; branched-chain amino acid;
KM corynebacterium; leucine; isoleucine; valine; medicine;
KM animal nutrition; ds.

Corynebacterium glutamicum.

PN EP1096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

AA 27-OCT-1999; 99DE-01051708.

PA (DEGS) DEGUSSA AG.

PA (KERT) FORSCHUNGSZENTRUM JOELICH GMBH.
XX

PI Kennerknecht N, Eggeling L, Sahm H, Pfeifferle W,
XY

DR WPI ; 2001-391595/42

XX

PT fermentative production of branched-chain amino acids.

PS Claim 5; Page 16; 23pp; German.

This invention describes a novel isolated polynucleotide (I) containing

CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC at least one sequence that (1) is /0% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)

CC polypeptides; (ii) encodes a polypeptide at least 10% identical with (3)
CC or (5); (iii) in the complement of (i) or (ii), or (iv) contains at least

CC 15 consecutive bases from (1) - (111). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4): (b)

CC coryneform microorganisms, especially *Corynebacterium*, transformed with

one or more (1), where these are replicative DNA; (c) production of branched-chain L-lea by fermentation of Corynebacterium bacteria in which the bme and/or bmf genes (or equivalent sequences) are amplified, and/or especially overexpressed; and (d) method for isolating the bme and/or bmf genes. (1) is used for transformation of Corynebacterium bacteria being used for fermentative production of branched-chain amino acids, specifically leucine, isoleucine and valine, which are useful in medicine and animal nutrition. (1) can also be used as source of primers and probes for isolation of related sequences. Transformation with (1) increases yield of branched-chain amino acids. This sequence encodes the Corynebacterium glutamicum ATCC 14752 bme protein described in the method of the invention

Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match	25.5%	Score 324	DB 4	Length 324
Best Local Similarity	100.0%	Pred. No.	4.4e-93	
Matches 324	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```

Oy      853 ATGACACTGATTTCCTCTGATTTCTCCTGTGTTCGGAGTATGCGAGCATTAATCTTT 912
        |||||
Db      1  ATGCAACTGATTTCTCTGATTTCTCCTGTGTTCGGAGTATGCGAGCATTAATCTTT 60

```

Qy 913 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAGTGAATCACAATTGTGGGC 972

Db 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAGTGAATCACAATTGTGGGC 120

Qy 973 AAATGGCGATGTCAGCAGGAATCTTCCATTGTACCGCATCAACGTTCCG 103
| | | | |
Db 121 AAATGGCGATGTCAGCAGGAATCTTCCATTGTACCGCATCAACGTTCCG 180

Dy 1033 AGCATGGCATAAGTCTGAGACTCTAACCTTTGGTCATATGCCGTGGATTACAGTG 109
| | | | |
Db 181 AGCATGGCATAAGTCTGAGACTCTAACCTTTGGTCATATGCCGTGGATTACAGTG 240

Qy 1093 GTGGCGCATCTTCTTGGCGGTCCAGCAGCACTTGTTAGAGGTTGGCGCTGGCACCATCGTT 115

Db 241 GTGGCGCATCTTCTTGGCGGTCCAGCAGCACTTGTTAGAGGTTGGCGCTGGCACCATCGTT 300

QY	1153 TTGTGGACTGCGAATCTTTC 1176
Db	301 TTGTGGACTGCGAATCTTTC 324

RESULT 10

ID AAH65255 standard; DNA; 324 BP

AC AAH65255

DT 26-SEP-2001 (first entry)

MM	
DE	C glutamicum coding sequence fragment SEQ ID NO: 290

XX . Corvneform bacterium: amino acid synthesis: vitamin: saccharide: KW

organic acid synthesis; db

Corynebacterium glutamicum

PN EP1108790-A2

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688

PR 16-DEC-1999; 99JP-00377484

PR	03-AUG-2000; 2000JP-00280988
FN	07-AER-2000; 2000UE-00135102

PA (KYOW) KYOWA HAKKO KOGYO KK

XX Nakagawa S. Mizouchi H. Ar

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

QY 769 TCACAGTCAGGCGCTATTGGCGCGCTGCTGATCTTCTTGGGCTGTGACCATCGGTA 828
 DB 480 TCACAGTCAGGCGCTATTGGCGCGCTGCTGATCTTCTTGGGCTGTGACCATCGGTA 539
 QY 829 CTTCTTCTTGGGAAAGGCTGCTTAA 953
 DB 540 CTTCTTCTTGGGAAAGGCTGCTTAA 564

RESULT 7

AAF61693/C
 ID AAF61693 standard; DNA; 397 BP.

AC AAF61693;
 DT 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX Corynebacterium glutamicum.

OS EPI090993-A1.

PN 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

XX (DEBS) DEGUSA-HUELS AG.

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;

DR WPI; 2001-292927/31.

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.

XX Example 3; Page 16; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention

XX Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

QY Query Match 28.8%; Score 365.6; DB 4; Length 397;

DB Best Local Similarity 98.9%; Pred. No. 2.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGGAATCTAGCTCATATATTGCAACATAGCTTGAAGTGGCAAACTGGCAACA 68

DB 376 AATGGAATCTAGCTCATATATTGCAACATAGCTTGAAGTGGCAAACTGGCAACA 317

QY 69 AAACATCCGGCAATGTGTGATGATTTAGTGTGAAAAAGCAAGAGATTCAATCAA 128

DB 316 AAACATCCGGCAATGTGTGATGATTTAGTGTGAAAAAGCAAGAGATTCAATCAA 257

QY 129 GCCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATATTAAGTTATCGGCGCT 188

DB 256 GCCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATATTAAGTTATCGGCGCT 197
 QY 189 ACGAAATCGGCAAGGCTTAAAACTCCCTTGTGACAGTTTGGCATGTACCGATTG 248
 DB 196 ACGAAATCGGCAAGGCTTAAAACTCCCTTGTGACAGTTTGGCATGTACCGATTG 137
 QY 249 GTATTGCGTTTGGTCTCTTGGTTATTCAATAGCGTACGAATGTGTGGAGAGCCCACTGT 308
 DB 136 GTATTGCGTTTGGTCTCTTGGTTATTCAATAGCGTACGAATGTGTGGAGAGCCCACTGT 77
 QY 309 TTTCGGACCTGATTTTGGCGGCTCCACCGAAATGTGTGTATCGGCTCTGTGTGGCGG 368
 DB 76 TTTCGGACCTGATTTTGGCGGCTCCACCGAAATGTGTGTATCGGCTCTGTGTGGCGG 17
 QY 369 CAGCGCCCTGG 380
 DB 16 CAGCGCCCTGG 5

RESULT 8

AAF61695/C
 ID AAF61695 standard; DNA; 778 BP.

AC AAF61695;

DT 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX Corynebacterium glutamicum.

OS EPI090993-A1.

PN 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

XX (DEBS) DEGUSA-HUELS AG.

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;

DR WPI; 2001-292927/31.

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.

XX Example 3; Page 17; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention

XX Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;

QY Query Match 28.8%; Score 365.6; DB 4; Length 778;

DB Best Local Similarity 98.9%; Pred. No. 3.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGGAATCTAGCTCATATATTGCAACATAGCTTGAAGTGGCAAACTGGCAACA 68

XX 02-MAY-2001.
 XX 11-OCT-2000; 2000EP-00122057.
 XX 27-OCT-1999; 99DE-01051708.
 XX (DEGS) DEGUSSA AG.
 XX (KERU) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX Kemerkecht N, Eggeling L, Sahn H, Pfeifferle W;
 XX WPI: 2001-391595/42.
 XX P-PSDB; AAB66247.
 XX New export genes from coryneform bacteria, useful for increasing
 XX fermentative production of branched-chain amino acids.
 XX Claim 5; Page 14-15; 23pp; German.
 XX This invention describes a novel isolated polynucleotide (1) containing
 XX at least one sequence that (i) is 70% identical with a sequence that
 XX encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 XX polypeptides; (11) encodes a polypeptide at least 70% identical with (3)
 XX or (5); (111) is the complement of (1) or (11), or (11v) contains at least
 XX 15 consecutive bases from (1) - (111). The invention also describes (a)
 XX protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 XX coryneform microorganisms, especially Corynebacterium, transformed with
 XX one or more (1), where these are replicative DNA; (c) production of
 XX branched-chain L-aa by fermentation of coryneform bacteria in which the
 XX brnB and/or brnF genes (or equivalent sequences) are amplified
 XX especially overexpressed; and (d) method for isolating the brnB and/or
 XX brnF genes. (1) is used for transformation of coryneform bacteria being
 XX used for fermentative production of branched-chain amino acids,
 XX specifically leucine, isoleucine and valine, which are useful in medicine
 XX and animal nutrition. (1) can also be used as source of primers and
 XX probes for isolation of related sequences. Transformation with (1)
 XX increases yield of branched-chain amino acids. This sequence encodes the
 XX Corynebacterium glutamicum ATCC 14752 brnF protein described in the
 XX method of the invention
 XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 XX
 XX Query Match 59.2%; Score 753; DB 4; Length 753;
 XX Best Local Similarity 100.0%; Pred. No. 1.3e-230; Indels 0; Gaps 0;
 XX Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 101 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGTGCCATCCAGGCAAGCCCTG 160
 DB 1 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGTGCCATCCAGGCAAGCCCTG 60
 QY 161 GAACCAATGATTAAGTTATGCGGCTTACGAAATCGGGCAAGCTTAAACCTCCCTT 220
 DB 61 GAACCAATGATTAAGTTATGCGGCTTACGAAATCGGGCAAGCTTAAACCTCCCTT 120
 QY 221 GGTGAGAGTTGGGATGATACCCGATTTGGTATTTGGTCTCTTGGTATTCATATC 280
 DB 121 GGTGAGAGTTGGGATGATACCCGATTTGGTATTTGGTCTCTTGGTATTCATATC 180
 QY 281 GGTGAGAGTTGGGATGATACCCGATTTGGTATTTGGTCTCTTGGTATTCATATC 340
 DB 181 GGTGAGAGTTGGGATGATACCCGATTTGGTATTTGGTCTCTTGGTATTCATATC 240
 QY 341 ATGCTGATGATGCGCTGTTGGGCGAGGCGCCCTGCGGATCGCGTCAACCA 400
 DB 241 ATGCTGATGATGCGCTGTTGGGCGAGGCGCCCTGCGGATCGCGTCAACCA 300
 QY 401 TTGCTGTGAATCTCGGCAAGTATTCATATGCGTTTCAATCCGCTGATGTGTCAA 460
 DB 301 TTGCTGTGAATCTCGGCAAGTATTCATATGCGTTTCAATCCGCTGATGTGTCAA 360
 QY 461 AACCCATGCGCGTTTCAATTCGTTTGGGCTTATGAGAGGCTACGACGATCACT 520

DB 361 AACCCATGCGCGTTTCAATTCGTTTGGGCTTATGAGAGGCTACGACGATCACT 420
 QY 521 GGGGCAAGCCCGAGAGGTGTGCGGCTGAGGCACTTATTCATGCAATAGCGTTTCA 580
 DB 421 GGGGCAAGCCCGAGAGGTGTGCGGCTGAGGCACTTATTCATGCAATAGCGTTTCA 480
 QY 581 TCTTACTGGTATTCGAGGCTGTCTACCGGATGCGATTCGAGAGTTGATTCCTTTGA 640
 DB 481 TCTTACTGGTATTCGAGGCTGTCTACCGGATGCGATTCGAGAGTTGATTCCTTTGA 540
 QY 641 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTTGTCAAGCTGACTTTGATTCCTGC 700
 DB 541 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTTGTCAAGCTGACTTTGATTCCTGC 600
 QY 701 CGAACGAAAAGACAGATTCCTTCTGCTGCTGCGAGTTGAGTTACCATTCCTCTT 760
 DB 601 CGAACGAAAAGACAGATTCCTTCTGCTGCTGCGAGTTGAGTTACCATTCCTCTT 660
 QY 761 GTGTATTCGAGTCAAGGCGCTTATTTGGGCGCTGCTGATCTTGTGAGTCTGTGAC 820
 DB 661 GTGTATTCGAGTCAAGGCGCTTATTTGGGCGCTGCTGATCTTGTGAGTCTGTGAC 720
 QY 821 ATCCGCTACTTCTTCTTGGGAAAGCTGTCTAA 853
 DB 721 ATCCGCTACTTCTTCTTGGGAAAGCTGTCTAA 753
 RESULT 5
 AAH65254
 ID AAH65254 standard; DNA; 753 BP.
 XX
 XX AAH65254;
 AC
 XX 26-SEP-2001 (first entry)
 DT
 XX
 XX C glutamicum coding sequence fragment SEQ ID NO: 289.
 DE
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KM organic acid synthesis; ds.
 KW
 XX Corynebacterium glutamicum.
 OS
 XX
 XX EP1108790-A2.
 PN
 XX 20-JUN-2001.
 PD
 XX
 XX 18-DEC-2000; 2000EP-00127688.
 PF
 XX
 XX 16-DEC-1999; 99DP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Nakagawa S, Mizoguchi H, Ardo S, Hayaishi M, Ochiai K, Yokoi H;
 FI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 XX WPI: 2001-376931/40.
 DR P-PSDB; AAG90035.
 DR
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 FT
 XX
 XX Claim 8; SEQ ID NO 289; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 XX are useful for identifying the mutation point of a gene derived from a
 XX mutant of coryneform bacterium, measuring expression amount and analysing
 XX the expression profile or expression pattern of a gene derived from
 XX Coryneform bacterium, and identifying a homologue of a gene derived from
 XX coryneform bacterium. Coryneform bacteria are useful for producing amino

PD		20-JUN-2001.
XX		
PF	18-DEC-2000;	2000EP-00127688.
XX		
PR	16-DEC-1999;	9AUP-00377484.
PR	07-APR-2000;	2000JP-00159162.
XX	03-AUG-2000;	2000JP-00280988.
XX		
PA	(KYOW) KYOMA HAKKO KOGYO KK.	
PI	Nakagawa S., Mizuguchi H., Ando S., Hayashi M., Ochiai K., Yokoi H.;	
PI	Tateishi N., Senoh A., Ikeda M., Ozaki A;	
XX		
DR	WPI; 2001-376931/40.	
XX		
PT	Novel polynucleotides derived from Corynebacterium bacterie, for identifying	
PT	mutation point of a gene, measuring expression of a gene, analyzing	
PT	expression profile or pattern of a gene and identifying homologous gene.	
XX		
PS	Claim 7; SEQ ID NO 1; 246pp + Sequence listing; English.	
XX		
CC	The present invention provides a number of nucleotide and protein	
CC	sequences from the Coryneform bacterium Corynebacterium glutamicum. These	
CC	are useful for identifying the mutation point of a gene derived from a	
CC	mutant of coryneform bacterium, measuring expression amount and analysing	
CC	the expression profile or expression pattern of a gene derived from	
CC	Coryneform bacterium, and identifying a homologue of a gene derived from	
CC	coryneform bacterium. Coryneform bacteria are useful for producing amino	
CC	acids, nucleic acids, vitamins, saccharides and organic acids,	
CC	particularly L-lysine. The present sequence is a nucleic acid described	
CC	in the exemplification of the invention. Note: The sequence data for this	
CC	patent did not form part of the printed specification, but was obtained	
CC	in electronic format directly from the European Patent Office	
XX		
SQ	Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;	
	Query Match	
	Best Local Similarity 100.0%; Score 1271; DB 5; Length 349980;	
	Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCGGATCAATGAATCTAGCTTCATTATATTTGGACAATAGCCTAGTtaggtggcgaac 60	
DB	276729 GGGCATCAATGAATCTAGCTTCATAATTTGACAAATAGCCTAGTtaggtggcgaac 60	
QY	61 TGGAACAACAACTACCGGGCAATTTGTGTGATTTGTAGTGCAAAAAGCAAGAT 120	
DB	276789 TGGCAACAAACTACCGGGCAATTTGTGTGATTTGTAGTGCAAAAAGCAAGAT 120	
QY	121 TCATTCAAAGCTGAGGTGTGCCATTCAAAGCACCTTGAAACGATGATTAAGGTTA 180	
DB	276849 TCATTCAAAGCTGAGGTGTGCCATTCAAAGCACCTTGAAACGATGATTAAGGTTA 180	
QY	181 TCGGGGCACGAATTCGGCGCAAGTCTTAAAACTCCCTGGTGGAGGTTTGGGCATGTA 240	
DB	276909 TCGGGGCACGAATTCGGCGCAAGTCTTAAAACTCCCTGGTGGAGGTTTGGGCATGTA 240	
QY	241 CCCAGTGGTATGTCGTTGTCCTCTTGTATTCAATACGGCTACGATGTTGAGGACG 300	
DB	276969 CCCAGTGGTATGTCGTTGTCCTCTTGTATTCAATACGGCTACGATGTTGAGGACG 300	
QY	301 CCACATGTTTTCCGGCTCGAATTTCCGGGGCTCAACCGAAATGCTGATTCGGCTCGT 360	
DB	277029 CCACATGTTTTCCGGCTCGAATTTCCGGGGCTCAACCGAAATGCTGATTCGGCTCGT 360	
QY	361 TGTGGGGCAGAGCCCCCTGGGGCGCAATCGGCGTCAACACATTCGTGTGAATTCGGCCA 420	
DB	277089 TGTGGGGCAGAGCCCCCTGGGGCGCAATCGGCGTCAACATTCGTGTGAATTCGGCCA 420	
QY	421 CGAATTCATATCGTTTCATTCGCGCTGATGTCGTCGCAAAACCCCATTCGCGCTTCTA 480	
DB	277149 CGAATTCATATCGTTTCATTCGCGCTGATGTCGTCGCAAAACCCCATTCGCGCTTCTA 480	
QY	481 TTGGGTTTGGCGCTTATCGACGAAGCTACAGACTGCGGGCCAGGCCCGCAGGCTG 540	

Db	277209	TTGGTTTGGGGCTTAATGAGAGAGCGCTACGAGTACATGGGGCCAGGCCCGGAGGCTG	277268
Qy	541	GTGGGGGTGGCGACTTATCTCAATGCAAAATAGCGTTTACCTCTACTGGGTAATTCGGCGG	600
Db	277269	GTGGGGGTGGCGACTTATCTCAATGCAAAATAGCGTTTACCTCTACTGGGTAATTCGGCGG	277328
Qy	601	TCACACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAAATTAAGGGCCTCGAGTTCCG	660
Db	277329	TCACACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAAATTAAGGGCCTCGAGTTCCG	277388
Qy	661	CCCTTCTCTCTCTTTGTCAGCTTACCTTTGGATTCCGCCGAGACGAAAAAGCAGATCCC	720
Db	277389	CCCTTCTCTCTCTTTGTCAGCTTACCTTTGGATTCCGCCGAGACGAAAAAGCAGATCCC	277448
Qy	721	TTCCTGCTGCTGCGAGTTTGAAGCTTCACCATTCCTCTTGGTGAATTCACAGTCAAGC	780
Db	277449	TTCCTGCTGCTGCGAGTTTGAAGCTTCACCATTCCTCTTGGTGAATTCACAGTCAAGC	277508
Qy	781	CCATTTTGGGGGGCGCTGATCTTTTGGGTCGTGTACCAATCCGGTACTTCTTCTGGG	840
Db	277509	CCATTTTGGGGGGCGCTGATCTTTTGGGTCGTGTGTACCAATCCGGTACTTCTTCTGGG	277568
Qy	841	AAAGCTGCTAAATGACAACTGATTTCTCTGATATCTCTTGTGTGTGAGTATGTGCA	900
Db	277569	AAAGCTGCTAAATGACAACTGATTTCTCTGATATCTCTTGTGTGTGAGTATGTGCA	277628
Qy	901	GTGATTACTTTTGGCGCTCCGGGGCGGTTCGGTTCTTAAATCCCTTAAGCCCTTACGTAATCA	960
Db	277629	GTGATTACTTTTGGCGCTCCGGGGCGGTTCGGTTCTTAAATCCCTTAAGCCCTTACGTAATCA	277688
Qy	961	CAATTTGTGGGCAAAATGGCGCATGTGGAATGCGACAGAGAAATCCTTGGCATTTTGACCGCA	1020
Db	277689	CAATTTGTGGGCAAAATGGCGCATGTGGAATGCGACAGAGAAATCCTTGGCATTTTGACCGCA	277748
Qy	1021	TCAACTGTTTGGAGCAATGCGATATGATGTAAGACTCTTAACTTGGTCTCATTTGGCGGT	1080
Db	277749	TCAACTGTTTGGAGCAATGCGATATGATGTAAGACTCTTAACTTGGTCTCATTTGGCGGT	277808
Qy	1081	GCGATTACAGTGTGTGGCGCATCTTCTTGGCGGTGACGACACTTGTGAACGTTGGCGCT	1140
Db	277809	GCGATTACAGTGTGTGGCGCATCTTCTTGGCGGTGACGACACTTGTGTGAACGTTGGCGCT	277868
Qy	1141	GGGACATGCTTTTGTGTGGAATGGTGAATCTTTCTTAAATCTGATTAATTAACAAAT	1200
Db	277869	GGGACATGCTTTTGTGTGGAATGGTGAATCTTTCTTAAATCTGATTAATTAACAAAT	277928
Qy	1201	CGGCATGCTCTCATATTGAAAGGGAGTGGCGAATTTTTAAGAACCTTGAAGAGGCTTAAG	1260
Db	277929	CGGCATGCTCTCATATTGAAAGGGAGTGGCGAATTTTTAAGAACCTTGAAGAGGCTTAAG	277988
Qy	1261	CAGACAGCGCT 1271	
Db	277989	CAGACAGCGCT 277999	
RESULT 4			
AAH21110			
XX	AAH21110	standard; DNA; 753 BP.	
XX	AAH21110;		
XX	AC		
XX	DT	05-SEP-2001 (first entry)	
XX	DE		
XX	XX	C. glutamicum brnf DNA.	
XX	KV	L- amino acid production; brnf; brng; branched-chain amino acid;	
XX	KW	coryneform bacterium; leucine; isoleucine; valine; medicine;	
XX	KX	animal nutrition; ds.	
XX	OS	Corynebacterium glutamicum.	
XX	FN	EP1096010-A1.	

CC at least one sequence that (1) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (11) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (11) is the complement of (1) or (11), or (11) contains at least
 CC 15 consecutive bases from (1)-(11). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially *Corynebacterium*, transformed with
 CC one or more (1), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC bnf and/or bnf genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the bnf and/or
 CC bnf genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC *Corynebacterium glutamicum* ATCC 14752 bnf and bnf proteins described in
 CC the method of the invention
 CC
 XX
 SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 1271; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATGATGGAATCTAGCTTCAATATTTGCAATAGCCTAGTTGAGTGGCGAAG 60
 DB 1 GCGGATGATGGAATCTAGCTTCAATATTTGCAATAGCCTAGTTGAGTGGCGAAG 60
 QY 61 TGGCAACAAACTACCCGCGCAATGTTGATGATGATGATGATGATGATGATGATGAT 120
 DB 61 TGGCAACAAACTACCCGCGCAATGTTGATGATGATGATGATGATGATGATGATGAT 120
 QY 121 TCATTCAAGCTGAGAGTGTGCGATCCAGAGCCCTGAAACCGATGATGATGATGATGAT 180
 DB 121 TCATTCAAGCTGAGAGTGTGCGATCCAGAGCCCTGAAACCGATGATGATGATGATGAT 180
 QY 181 TCGGCGCTACGAAATCGCGCAAGTGTGATGATGATGATGATGATGATGATGATGAT 240
 DB 181 TCGGCGCTACGAAATCGCGCAAGTGTGATGATGATGATGATGATGATGATGATGAT 240
 QY 241 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 301 CCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 420
 DB 361 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 420
 QY 421 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 480
 DB 421 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 480
 QY 481 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 540
 DB 481 TGTGGGCGCAAGCCCTGCGCGCATGCGGCTGACGACATGCTGTGATGATGATGATGAT 540
 QY 541 GTGGCGGCGGAGCTTATCTCAATGCAATAGCCTTCACTCTAGCGGATGATGATGAT 600
 DB 541 GTGGCGGCGGAGCTTATCTCAATGCAATAGCCTTCACTCTAGCGGATGATGATGAT 600
 QY 601 TCTACCGGAGTGGCGATGCGAGGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 601 TCTACCGGAGTGGCGATGCGAGGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 661 CATTGCTCTCTCTTTGCAAGCTGACTTGGATCTCTGCGCAAGCAAAAAGCAGATGCC 720
 DB 661 CATTGCTCTCTCTTTGCAAGCTGACTTGGATCTCTGCGCAAGCAAAAAGCAGATGCC 720

QY 721 TTCTCTGCTGCTGCGAGGTTTGAAGCTTCACTGCTCTTGTGATTTCCAGGTCAGGC 780
 DB 721 TTCTCTGCTGCTGCGAGGTTTGAAGCTTCACTGCTCTTGTGATTTCCAGGTCAGGC 780
 QY 781 CATTGTCGCGCGCTGCTGATCTTCTTGGAGTGTGACCATCCGATCTTCTTGGG 840
 DB 781 CATTGTCGCGCGCTGCTGATCTTCTTGGAGTGTGACCATCCGATCTTCTTGGG 840
 QY 841 AAGGCTGCTAAATGCACTGATTTCTCTGATTTCTCTGATTTCTGTCAGTATGCA 900
 DB 841 AAGGCTGCTAAATGCACTGATTTCTCTGATTTCTCTGATTTCTGTCAGTATGCA 900
 QY 901 GTGATTAATTTTGGCTGCGCGGATCCGTTCTTAATCTTAAGCCCTAAGTATCA 960
 DB 901 GTGATTAATTTTGGCTGCGCGGATCCGTTCTTAATCTTAAGCCCTAAGTATCA 960
 QY 961 CAATTGTCGCGCAAAATGCGAGTGTGATGATGATGATGATGATGATGATGATGAT 1020
 DB 961 CAATTGTCGCGCAAAATGCGAGTGTGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 TCAAGCTTTGCGAGCATGCGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 TCAAGCTTTGCGAGCATGCGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 GCGATTACAGTGTGCGCGCATCTTCTTGGCGGTCGACGACCTTGTGAGCGTCT 1140
 DB 1081 GCGATTACAGTGTGCGCGCATCTTCTTGGCGGTCGACGACCTTGTGAGCGTCT 1140
 QY 1141 GGCACATCGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 GGCACATCGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTAAGAAAGGCTTAA 1260
 DB 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTAAGAAAGGCTTAA 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 2
 AAH21112
 ID AAH21112 standard; DNA; 1271 BP.
 AC AAH21112;
 DT 05-SEP-2001 (first entry)
 XX
 C. glutamicum DNA encoding bnf and bnf.
 L-amino acid production; bnf; bnf; branched-chain amino acid;
 coryneform bacterium; leucine; isoleucine; valine; medicine;
 animal nutrition; ds.
 OS
 XX
 Corynebacterium glutamicum.
 XX
 Key location/Qualifiers
 FH 101..856
 FT /tag= a
 FT /product= "bnf"
 FT 853..1179
 FT /tag= b
 FT /product= "bnf"
 XX
 EP1096010-A1.
 XX
 PD 02-MAY-2001.
 XX
 PF 11-OCT-2000; 2000EP-00122057.
 XX
 PR 27-OCT-1999; 99DE-01051706.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: April 19, 2004, 10:56:53 ; Search time 591.9 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-6
Perfect score: 1271
Sequence: 1 ggcgcgacaaaggaatctag.....agccttaagcagacagcgc 1271

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	AAH21109	Aah21109 C. glutam
2	1271	100.0	1271	AAH21112	Aah21112 C. glutam
3	1271	100.0	349980	AAH64966	Aah64966 C. glutam
4	753	59.2	753	AAH21110	Aah21110 C. glutam
5	753	59.2	753	AAH65254	Aah65254 C. glutam
6	553	43.5	564	AAH61969	Aah61969 C. glutam
7	365.6	28.8	397	AAH61693	Aah61693 C. glutam
8	365.6	28.8	778	AAH61695	Aah61695 C. glutam
9	324	25.5	324	AAH21111	Aah21111 C. glutam
10	324	25.5	324	AAH65255	Aah65255 C. glutam
11	324	25.5	324	AAH61970	Aah61970 C. glutam
12	185	14.6	715	AAH61688	Aah61688 C. glutam
13	177	13.9	177	AAH61968	Aah61968 C. glutam
14	127.4	10.0	711	AAH68373	Aah68373 C. glutam
15	127.4	10.0	711	AAH61294	Aah61294 C. glutam
16	127.4	10.0	309400	AAH68534	Aah68534 C. glutam
17	48.2	3.8	2000	ADAV1938	Adav1938 Rice gene
18	48	3.8	10732	AAH610594	Aah610594 Gene enco
19	40.8	3.2	7645	AAH659610	Aah659610 Human gen
20	40.2	3.2	4947	AAH68927	Aah68927 TSPs17 po
21	39.8	3.1	768	ADCC91639	Adcc91639 E. faeciu
22	39.2	3.1	6741	AAH10595	Aah10595 Gene enco
23	38.2	3.0	946	AAH55683	Aah55683 Human bre

ALIGNMENTS

24	37.6	3.0	759	7	ACF69125	ACF69125 Photorhab
25	37.6	3.0	5059	2	AAH84332	Aah84332 Steath v
26	37.6	3.0	110000	7	AAH67367_20	Continuation (21 o
27	37.6	3.0	243072	7	ACF65382	ACF65382 Photorhab
28	37.2	2.9	1024	6	ABX66366	Abx66366 Helicobac
29	37.2	2.9	2000	7	ADA71938	Ada71938 Rice gene
30	36.8	2.9	1230	7	ACA442205	ACA442205 Prokaryot
31	36.4	2.9	751	2	AAV58489	AAV58489 3' fragme
32	36.4	2.9	751	3	AAV61146	AAV61146 3' CDNA s
33	36.4	2.9	751	3	AAH06252	AAH06252 Human imm
34	36.4	2.9	751	3	ABH71157	Abh71157 Human pro
35	36.4	2.9	751	4	AAH93368	Aah93368 Human pro
36	36.4	2.9	751	4	AAH63460	Aah63460 Human pro
37	36.4	2.9	751	4	AAH02433	Aah02433 Prostate
38	36.4	2.9	751	4	AAH84682	Aah84682 Human pro
39	36.4	2.9	751	5	ACA59269	ACA59269 Prostate
40	36.4	2.9	751	5	AAH10011	AAH10011 Human pro
41	36.4	2.9	751	6	ABH34832	Abh34832 Human NI-
42	36.4	2.9	751	6	ABH58541	Abh58541 Prostate
43	36.4	2.9	751	7	ACC94996	Acc94996 Prostate
44	36.4	2.9	751	9	ADB13462	Adb13462 Human pro
45	36.2	2.8	19012	2	AAQ36952	Aaq36952 HSA gene.

AAH21109	AAH21109 standard; DNA; 1271 BP.
AAH21109;	
05-SEP-2001	(first entry)
C. glutamicum DNA encoding brnF and brnE.	
L-amino acid production; brnF; brnE; branched-chain amino acid;	
Coryneform bacterium; leucine; isoleucine; valine; medicine;	
animal nutrition; ds.	
Corynebacterium glutamicum.	
Key	Location/Qualifiers
CDS	101..856
FT	/*tag= a "brnF"
FT	/product= "brnF"
FT	853..1179
FT	/*tag= b
FT	/product= "brnE"
EP1096010-A1.	
02-MAY-2001.	
11-OCT-2000; 2000EP-00122057.	
27-OCT-1999; 99DS-01051708.	
(DEGS) DEGUSSA AG.	
(KERU) FORSCHUNGSZENTRUM JUELICH GMBH.	
Kernerkrecht N, Eggeling L, Sahn H, Pfeifferle W;	
WPI; 2001-391595/42.	
P-PSDB; AAB86247, AAB86248.	
New export genes from coryneform bacteria, useful for increasing	
fermentative production of branched-chain amino acids.	
Claim 4 (1); Page 13; 23pp; German.	
This invention describes a novel isolated polynucleotide (I) containing	


```

Db      181 GGCCTAGATGATGGGCGACCCCACTGTTTCGCGCTGATTTTCGCGGCTCCACCGAA 240
Qy      341 ATGCTGATCATGCGCTCGTGTGTGGGCGAGGCGCCCTGGGCGCATGCGCTCACCACA 400
Db      241 ATGCTGATCATGCGCTCGTGTGTGGGCGAGGCGCCCTGGGCGCATGCGCTCACCACA 300
Qy      401 TTGCTGATGAACTTCCGCCACGATTCATGCGTTTCATTCGCCGTCATGCTGCAAA 460
Db      301 TTGCTGATGAACTTCCGCCACGATTCATGCGTTTCATTCGCCGTCATGCTGCAAA 360
Qy      461 AACCCATGCGCGCTTCTATTCGTTTTCGCGCTTATGACGAAAGCTTACGACT 520
Db      361 AACCCATGCGCGCTTCTATTCGTTTTCGCGCTTATGACGAAAGCTTACGACT 420
Qy      521 GCGGCGAGGCGCGGCTGATGCGGCTGCGGCTTATGATCAATGCAATAGCTTTCAC 580
Db      421 GCGGCGAGGCGCGGCTGATGCGGCTGCGGCTTATGATCAATGCAATAGCTTTCAC 480
Qy      581 TCCTACTGGGTATTCGCGGCTTCACCGAGTGGCATGCGAGTTGATTCCTTTGAA 640
Db      481 TCCTACTGGGTATTCGCGGCTTCACCGAGTGGCATGCGAGTTGATTCCTTTGAA 540
Qy      641 ATTAAGGCTCGAGTTTCGCGCTTCTCTCTCTTTCGACGCTGACTTGTGATTCCTG 700
Db      541 ATTAAGGCTCGAGTTTCGCGCTTCTCTCTCTTTCGACGCTGACTTGTGATTCCTG 600
Qy      701 CGAAGCAAAAGAGATCCCTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATGCTCT 760
Db      601 CGAAGCAAAAGAGATCCCTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATGCTCT 660
Qy      761 GTGCTATTCAGGTCAGGCTTATTTGGCGGCTGCTGATCTTCTGGGTCGTGACC 820
Db      661 GTGCTATTCAGGTCAGGCTTATTTGGCGGCTGCTGATCTTCTGGGTCGTGACC 720
Qy      821 ATCCGACTTCTCTTGGGAAAGGCTGCTAAA 853
Db      721 ATCCGACTTCTCTTGGGAAAGGCTGCTAAA 753

```

```

RESULT 15
AX137083/c 397 bp DNA linear PAT 30-MAY-2001
LOCUS      AX137083
DEFINITION Sequence 7 from Patent EP1090933.
ACCESSION AX137083
VERSION    AX137083.1 GI:14273428
KEYWORDS
SOURCE
ORGANISM   Corynebacterium glutamicum
            Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE  1 Moeckel,B., Pfefferle,W., Puehler,A., Kallnowski,J. and Barhe,B.
            Nucleotide sequences coding for the lrp gene
            Patent: EP 1090933-A 7 11-Apr-2001;
            Degussa AG (DE)
FEATURES
            source
            1..397
            location/Qualifiers
            /organism="Corynebacterium glutamicum"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1718"
            /note="Lrp-Tail 1"

```

```

ORIGIN
Query Match      28.8%; Score 365.6; DB 6; Length 397;
Best Local Similarity 98.9%; Pred. No. 2e-89;
Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy      9 AATGGAATCTAGCTTATATTTGCAATAAGCCCTAGTGAAGTGGCGAAACTGGCAACA 68
Db      376 AATGGAATCTAGCTTATATTTGCAATAAGCCCTAGTGAAGTGGCGAAACTGGCAACA 317
Qy      69 AAACCTACCGGCAATTGTGTGATGATTTAGTGTGCAAAAAACGAAAGATTCATTCPA 128

```

```

Db      316 AAACCTACCGGCAATTGTGTGATGATTTAGTGTGCAAAAAACGAAAGATTCATTCPA 257
Qy      129 GCTGAGAGTGTGAGCATCCAGGAGGAGCCCTGGAAACGATGATTAAGTTATCGCGCT 188
Db      256 GCTGAGAGTGTGAGCATCCAGGAGGAGCCCTGGAAACGATGATTAAGTTATCGCGCT 197
Qy      189 ACGAATCGCGCAAGGCTTAATAAAGCTCTGCTGCAAGTTTGGCATGTACCGATTG 248
Db      196 ACGAATCGCGCAAGGCTTAATAAAGCTCTGCTGCAAGTTTGGCATGTACCGATTG 137
Qy      249 GTATTGGTTTGGTCTGCTGCTGCTTATTCATATACGAGTACGATGATGAGGAGGAGG 308
Db      136 GTATTGGTTTGGTCTGCTGCTGCTTATTCATATACGAGTACGATGATGAGGAGGAGG 77
Qy      309 TTTCCGAGCTGATTTTTCGCGGCTCCACCGAAATGCTGTCATGCGCTTGTGTGAGCG 368
Db      76 TTTCCGAGCTGATTTTTCGCGGCTCCACCGAAATGCTGTCATGCGCTTGTGTGAGCG 17
Qy      369 CAGCGCCCTGCG 380
Db      16 CAGCGCCCTGCG 5

```

Search completed: April 19, 2004, 18:30:05
Job time : 5391 secs


```

REFERENCE 1 (bases 1 to 753)
AUTHORS KernerKunehito,N., Sahn,H., Eggering,U. and Pfeifferle,W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
        method of isolating the same and utilization thereof
JOURNAL Patent: JP 200169788-A 2 26-JUN-2001;
        DEBUSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC14752
        PD JP 200169788-A/2
        PD 26-JUN-2001
        PR 24-OCT-2000 JP 2000324315
        PR 27-OCT-1999 DE 19951708.8
        PI NICOLE KERNERKUNEHITO,HERMANN SAHM,LOTHAR EGGERING,WALTER PI
FEATURES
source 1..753
        Location/Qualifiers
            1..753
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 GTGCAAAAAGCCAGAGATTCATTCAGCCCTGAGGTGCGCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAAGCCAGAGATTCATTCAGCCCTGAGGTGCGCATCCAGGAGCCCTG 60
QY 161 GAACAGATGATAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 220
DB 61 GAACAGATGATAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 120
QY 221 GGTGAGGTTTGCGCATGTACCCCATTTGCTTGTCTCTTGGTTATTCATAC 280
DB 121 GGTGAGGTTTGCGCATGTACCCCATTTGCTTGTCTCTTGGTTATTCATAC 180
QY 281 GGTACGATGATGAGGAGCCAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACGAA 340
DB 181 GGTACGATGATGAGGAGCCAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACGAA 240
QY 341 ATGCTGATGATGCGCTGTTGTTGCGGAGCGAGCCCTGCGCATCGCGTCCACCA 400
DB 241 ATGCTGATGATGCGCTGTTGTTGCGGAGCGAGCCCTGCGCATCGCGTCCACCA 300
QY 401 TTGCTGATGATTCGCGCAGCATTCATGCGTTTCAATCCGCTGATGTGTCAAA 460
DB 301 TTGCTGATGATTCGCGCAGCATTCATGCGTTTCAATCCGCTGATGTGTCAAA 360
QY 461 AACCCCATGCGCGTTTCAATTCGTTTTCGCGTTATGACGAAAGCTTCAAGCTACT 520
DB 361 AACCCCATGCGCGTTTCAATTCGTTTTCGCGTTATGACGAAAGCTTCAAGCTACT 420
QY 521 GCGGCGAGGCGCGAGGCTGTCGCGGTCGCGCATTCATTCGAAATGAAATAGGTTTAC 580
DB 421 GCGGCGAGGCGCGAGGCTGTCGCGGTCGCGCATTCATTCGAAATGAAATAGGTTTAC 480
QY 581 TCTCTATGAGTATTCGCGGCTTCACCGAGTGGCGCATCGCAGATTGATTCCTTTGAA 640
DB 481 TCTCTATGAGTATTCGCGGCTTCACCGAGTGGCGCATCGCAGATTGATTCCTTTGAA 540
QY 641 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTTGACAGCTGACTTGAATTCCTG 700
DB 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTTGACAGCTGACTTGAATTCCTG 600
QY 701 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCAATTCCTTCT 760

```

```

DB 601 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCAATTCCTTCT 660
QY 761 GTGTATATTCAGAGTCAGGCGCTATTTGCGGCTGTCGATCTTGTGGTCTGTTAAC 820
DB 661 GTGTATATTCAGAGTCAGGCGCTATTTGCGGCTGTCGATCTTGTGGTCTGTTAAC 720
QY 821 ATCCGATCTTCTTCTTGGGAAAGGCTGCTAAA 853
DB 721 ATCCGATCTTCTTCTTGGGAAAGGCTGCTAAA 753
RESULT 14
BD162490
LOCUS BD162490 753 bp DNA linear PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION BD162490.1 GI:27868248
VERSION JP 2002191370-A/289.
KEYWORDS unidentifed
SOURCE unidentifed
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 753)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
        Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotide
JOURNAL Patent: JP 2002191370-A 289 09-JUN-2002;
        KYOMA HAKKO KOGYO CO LTD
COMMENT OS Corynebacterium glutamicum
        PN JP 2002191370-A/289
        PD 09-JUN-2002
        PR 15-DEC-2000 JP 2000405096
        PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
        KEIICHI OCHIAI,
        HARUHIKO YOKOI,NAOKO TATEISHI,AKIHINO SENOO,MASATO IKEDA,AKIO
        PI OZAKI
        PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
        C12N1/15,
        PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
        PC 04,C12P13/08,
        PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
        G01N33/566
        PC G01N33/569,G01N33/68,G01N37/00/C12P21/08,(C12N1/21,C12R1:15),
        PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
        PC C12N15/00,C12N15/00
        CC Novel polynucleotide
        CC Novel polynucleotide
        FH Key
        FT source 1..753
        Location/Qualifiers
            1..753
            /organism="Corynebacterium glutamicum".
            Location/Qualifiers
            1..753
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 GTGCAAAAAGCCAGAGATTCATTCAGCCCTGAGGTGCGCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAAGCCAGAGATTCATTCAGCCCTGAGGTGCGCATCCAGGAGCCCTG 60
QY 161 GAACAGATGATAAGATTACGCGCTACGAATCGCGCAAGGTCTAAAACTCCCTT 220
DB 61 GAACAGATGATAAGATTACGCGCTACGAATTCGCGCAAGGTCTAAAACTCCCTT 120
QY 221 GGTGAGGTTTGCGCATGTACCCCATTTGCTTGTCTCTTGGTTATTCATAC 280
DB 121 GGTGAGGTTTGCGCATGTACCCCATTTGCTTGTCTCTTGGTTATTCATAC 180
QY 281 GGTACGATGATGAGGAGCCAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACGAA 340

```

QY 281 GGCTACGAATGGTGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 340
 DB 181 GGCTACGAATGGTGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 240
 QY 341 ATGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 400
 DB 241 ATGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 300
 QY 401 TTGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 460
 DB 301 TTGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 360
 QY 461 AACCCATTGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 520
 DB 361 AACCCATTGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 420
 QY 521 GCGGCGAGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 580
 DB 421 GCGGCGAGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 480
 QY 581 TCCTACTGAGGATTCGCGCGGCTCAGCGAGTGGCGAGTGGATTCCTTTGAA 640
 DB 481 TCCTACTGAGGATTCGCGCGGCTCAGCGAGTGGCGAGTGGATTCCTTTGAA 540
 QY 641 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTTGATTCCTGC 700
 DB 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTTGATTCCTGC 600
 QY 701 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCACATTCCTCT 760
 DB 601 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCACATTCCTCT 660
 QY 761 GTGATTAATTCAGATCGAGCCCTATTGCGCGCGCTGATGCTTCTTGGGCTGTTGAC 820
 DB 661 GTGATTAATTCAGATCGAGCCCTATTGCGCGCGCTGATGCTTCTTGGGCTGTTGAC 720
 QY 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 853
 DB 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 12
 AX137710 753 bp DNA linear PART 30-MAY-2001
 LOCUS AX137710
 DEFINITION Sequence 2 from Patent EP1096010.
 ACCESSION AX137710
 VERSION AX137710.1 GI:14273889
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriidae; Corynebacterium;
 1
 Kennerknecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 2 02-MAY-2001;
 Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 FEATURES
 source
 1..753
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1..>753
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39925.1"
 /db_xref="GI:14273890"
 /db_xref="REBASE:CA39925"
 /translation="MOTQRIHSLSEVSPKALPEDDKGYRREYIAQGLKSLAAGL
 GMPIGAFELVIVGYEWMAPIFSGILFASSTEMVIVLVGAAPGAYALTTLL

CDS
 1..753
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39925.1"
 /db_xref="GI:14273890"
 /db_xref="REBASE:CA39925"
 /translation="MOTQRIHSLSEVSPKALPEDDKGYRREYIAQGLKSLAAGL
 GMPIGAFELVIVGYEWMAPIFSGILFASSTEMVIVLVGAAPGAYALTTLL

ORIGIN

Query Match 59.2%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 3.3e-196;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

VNRHVFVAFSEPLHVKNPIARFYSVFALIDEAVATAPAGNSAMWLSMOIAFH
 SWVFGALGVIAELIPEIKGELFALFVTLTLDSCRTKQIPSLILAGLSFTI
 ALVIRGQALFALLIFLLLTIRYFELKAK"

QY 101 GTGCAAAAAGAGATTCATTCAGCCCTGAGAGTGTGCGCATCGAGGAGCCCTG 160
 DB 1 GTGCAAAAAGAGATTCATTCAGCCCTGAGAGTGTGCGCATCGAGGAGCCCTG 60
 QY 161 GAACCAATGATTAAGGTATCGCGCTACGAAATCGCGCAAGTCTAAACCTTCCTT 220
 DB 61 GAACCAATGATTAAGGTATCGCGCTACGAAATCGCGCAAGTCTAAACCTTCCTT 120
 QY 221 GCTGAGGTTTGGGCAATGATCCGATGATGATGCGTTGCTGCTGTTGATTAATAC 280
 DB 121 GCTGAGGTTTGGGCAATGATCCGATGATGATGCGTTGCTGCTGTTGATTAATAC 180
 QY 281 GGCTACGAATGGTGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 340
 DB 181 GGCTACGAATGGTGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGGCTCCACCGAA 240
 QY 341 ATGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 400
 DB 241 ATGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 300
 QY 401 TTGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 460
 DB 301 TTGCTGTCATCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 360
 QY 461 AACCCATTGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 520
 DB 361 AACCCATTGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 420
 QY 521 GCGGCGAGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 580
 DB 421 GCGGCGAGCGCCCTGTTTGGGCGAGCGCCCTGGGGCCATCGCGTCAACCA 480
 QY 581 TCCTACTGAGGATTCGCGCGGCTCAGCGAGTGGCGAGTGGATTCCTTTGAA 640
 DB 481 TCCTACTGAGGATTCGCGCGGCTCAGCGAGTGGCGAGTGGATTCCTTTGAA 540
 QY 641 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTTGATTCCTGC 700
 DB 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTTGATTCCTGC 600
 QY 701 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCACATTCCTCT 760
 DB 601 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGTTTGAAGTTCACATTCCTCT 660
 QY 761 GTGATTAATTCAGATCGAGCCCTATTGCGCGCGCTGATGCTTCTTGGGCTGTTGAC 820
 DB 661 GTGATTAATTCAGATCGAGCCCTATTGCGCGCGCTGATGCTTCTTGGGCTGTTGAC 720
 QY 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 853
 DB 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 13
 BD014992 753 bp DNA linear PART 27-AUG-2002
 LOCUS BD014992
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014992
 VERSION BD014992.1 GI:22555799
 KEYWORDS
 SOURCE unclassified
 ORGANISM unclassified

Db 663 GTCATTACTTTGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTGATCA 604
Qy 961 CAATTTGGGCAAAATATGGGATGTGATGCCAGAGAAATCCCTTGCATTTTACCGCA 1020
Db 603 CAATTTGGGCAAAATATGGGATGTGATGCCAGAGAAATCCTTGCATTTTACCGCA 544
Qy 1021 TCAACGTTTGGCAGCAATGCGATGATCTGAAGCTTAACCTTTGGTCTCATTTGGCGTT 1080
Db 543 TCAACGTTTGGCAGCAATGCGATGATCTGAAGCTTAACCTTTGGTCTCATTTGGCGTT 484
Qy 1081 GCGATTACAGTGTGGGCGCATCTTCTTGGCGGTGACGCACTTGTGTAGCGGTGGCGCT 1140
Db 483 GCGATTACAGTGTGGGCGCATCTTCTTGGCGGTGACGCACTTGTGTAGCGGTGGCGCT 424
Qy 1141 GCGACCATCGTTTGTGTGAGCTGTGATCTTTTCTTAAACTGCAATTAACAAAT 1200
Db 423 GCGACCATCGTTTGTGTGAGCTGTGATCTTTTCTTAAACTGCAATTAACAAAT 364
Qy 1201 CCGCATGCCCTCAATTTGAAGGGATGCGGATTTTAAAGAACTTAA-AAAAGCTTAA 1259
Db 363 CCGCATGCCCTCAATTTGAAGGGATGCGGATTTTAAAGAACTTAAAGAAAGCTTAA 304
Qy 1260 GCAAGCAGCGCT 1271
Db 303 GCAAGCAGCGCT 292

RESULT 10
LOCUS AR391954 753 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6613545.
ACCESSION AR391954
VERSION AR391954.1 GI:40115725
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 753)
TITLE Kemerinecht,N., Salm,H., Eggeling,U. and Pfeifferle,W.
Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 2 02-SEP-2003;
FEATURES
Location/Qualifiers
1..753
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

Qy 101 GTGCAAAAAACGCAAGATTCATTCAGGCTGAGGTGTGCGCCATCCAGAGAGCCCTG 160
Db 1 GTGCAAAAAACGCAAGATTCATTCAGGCTGAGGTGTGCGCCATCCAGAGAGCCCTG 60

Qy 161 GAACCAAGATGATTAAGTTATGCGGCTAAGAAATCGGCAAGGTCTAAAAACCTCCCTT 220
Db 61 GAACCAAGATGATTAAGTTATGCGGCTAAGAAATCGGCAAGGTCTAAAAACCTCCCTT 120

Qy 221 GCTGCAAGTTTGGGATGTACCGCATTTGGTATTTGGTGTCTCTTGGTATTCATATAC 280
Db 121 GCTGCAAGTTTGGGATGTACCGCATTTGGTATTTGGTGTCTCTTGGTATTCATATAC 180

Qy 281 GGTACAGATGATGTGGGCAAGCTGCTTTTCCGCGCTGATTTTGGCGGGCTCCAGCA 340
Db 181 GGTACAGATGATGTGGGCAAGCTGCTTTTCCGCGCTGATTTTGGCGGGCTCCAGCA 240

Qy 341 ATGCTGATCATGCGCCCTGTTGTGGGCGAGGCGCCCTGAGGCGCATGCGCTCACCA 400
Db 241 ATGCTGATCATGCGCCCTGTTGTGGGCGAGGCGCCCTGAGGCGCATGCGCTCACCA 300

Qy 401 TTGCTGTGAACCTTCCGCAAGATTCATGCGTTTCAATCCGCTGCATGTGTGCA 460
Db 301 TTGCTGTGAACCTTCCGCAAGATTCATGCGTTTCAATCCGCTGCATGTGTGCA 360

Qy 461 AACCCCATTTGCCGTTTCTATTGCGTTTCCGCTTATGCAAGAAAGCTTACGAGTCACT 520
Db 361 AACCCCATTTGCCGTTTCTATTGCGTTTCCGCTTATGCAAGAAAGCTTACGAGTCACT 420

Qy 521 GGGGCAAGGCGGAGGCGGTGCGGCTGAGCACTTATCTGAATGCAATAGCGTTTAC 580
Db 421 GGGGCAAGGCGGAGGCGGTGCGGCTGAGCACTTATCTGAATGCAATAGCGTTTAC 480

Qy 581 TCTTAAGTGTGATTTGGGCGGTCTCAACGAGTGTGCAATGCAAGTGTATCTTTGAA 640
Db 481 TCTTAAGTGTGATTTGGGCGGTCTCAACGAGTGTGCAATGCAAGTGTATCTTTGAA 540

Qy 641 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGCAAGCTTGTGATCTGCG 700
Db 541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGCAAGCTTGTGATCTGCG 600

Qy 701 CGAACGAAAAACAGATCCCTTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTT 760
Db 601 CGAACGAAAAACAGATCCCTTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTT 660

Qy 761 GTGTATTTCCAGTTCAGGCGCTTATTTGGGCGGTGATGTTGTTGGTGTGTGACC 820
Db 661 GTGTATTTCCAGTTCAGGCGCTTATTTGGGCGGTGATGTTGTTGGTGTGTGACC 720

Qy 821 ATCCGTTACTTCTTCTTGGGAAAGCTGTAA 853
Db 721 ATCCGTTACTTCTTCTTGGGAAAGCTGTAA 753

RESULT 11
LOCUS AX120373 753 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 289 from Patent EP1108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037086
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 289 20-JUN-2001;
KYOMA HAKKO KOGYO CO., LTD. (JP)
FEATURES
Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

Qy 101 GTGCAAAAAACGCAAGATTCATTCAGGCTTGAAGTGTGCGCATCCAGAGAGCCCTG 160
Db 1 GTGCAAAAAACGCAAGATTCATTCAGGCTTGAAGTGTGCGCATCCAGAGAGCCCTG 60

Qy 161 GAACCAAGATGATTAAGTTATGCGGCTAAGAAATCGGCAAGGTCTAAAAACCTCCCTT 220
Db 61 GAACCAAGATGATTAAGTTATGCGGCTAAGAAATCGGCAAGGTCTAAAAACCTCCCTT 120

Qy 221 GCTGCAAGTTTGGGATGTACCGCATTTGGTATTTGGTGTCTCTTGGTATTCATATAC 280
Db 121 GCTGCAAGTTTGGGATGTACCGCATTTGGTATTTGGTGTCTCTTGGTATTCATATAC 180

QY 841 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 900
 Db 277569 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 277628
 QY 901 GTGATTAATTTTCCGCTCCGCGGCGTTCCGTTCTTAATCTTAACCCCTACGTGAATCA 960
 Db 277629 GTGATTAATTTTCCGCTCCGCGGCGTTCCGTTCTTAATCTTAACCCCTACGTGAATCA 277688
 QY 961 CAATTGTGGGCAAAATGCGATGTGATGCGACGAGAAATCTTGCAATTTTGAACGCA 1020
 Db 277689 CAATTGTGGGCAAAATGCGATGTGATGCGACGAGAAATCTTGCAATTTTGAACGCA 277748
 QY 1021 TCAAGCTTTGCGAGCAATGAGATGATGAGAACTCTTAACCTTTGTTCTATGCGCTT 1080
 Db 277749 TCAAGCTTTGCGAGCAATGAGATGATGAGAACTCTTAACCTTTGTTCTATGCGCTT 277808
 QY 1081 GCGATTACAGTGTGCGGCACTCTTCTGCGGTCGACGCACTTGTGAGCGTTGCGGCT 1140
 Db 277809 GCGATTACAGTGTGCGGCACTCTTCTGCGGTCGACGCACTTGTGAGCGTTGCGGCT 277868
 QY 1141 GGCACCATGTTTTTGTGGAATGATCTTTTCTAAACTGATTAATTAACAAAT 1200
 Db 277869 GGCACCATGTTTTTGTGGAATGATCTTTTCTAAACTGATTAATTAACAAAT 277928
 QY 1201 CCGCATGCTCTCAATTTGAAGGGATGCGAATTTTGAAGAACTGAAGAAAGCTTAAG 1260
 Db 277929 CCGCATGCTCTCAATTTGAAGGGATGCGAATTTTGAAGAACTGAAGAAAGCTTAAG 277988
 QY 1261 CAGACAGCGCT 1271
 Db 277989 CAGACAGCGCT 277999

RESULT 8
 AX120085 349980 bp DNA linear PAT 11-MAY-2001
 LOCUS AX120085
 DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:14036800
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 Corynebacteriaceae; Corynebacterinae.
 REFERENCE 1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Akihara, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
 Novel polynucleotides
 Patent: EP 1108790-A 1 20-JUN-2001;
 JOURNAL KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source location/Qualifiers
 1. 349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0.000.001 0.349.980 349.980"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 349980;
 Best Local Similarity 100.0%; Fred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 276849 TCATTCAAGCTGAGAGTGTCCCATCCAGGAGACGCTTGAACCAATGATTAAGGTTA 276908
 QY 181 TCGGCGCTAGCAAAATCGGCGAAGCTTAATAAACCCTTGTGCGAGTTGGGATGTA 240
 Db 276909 TCGGCGCTAGCAAAATCGGCGAAGCTTAATAAACCCTTGTGCGAGTTGGGATGTA 276968
 QY 241 CCGGATGATTTGCGTTGTTGTTCTTGTGTTATTAATACGCTACGAATGTTGGGAGC 300
 Db 276969 CCGGATGATTTGCGTTGTTGTTCTTGTGTTATTAATACGCTACGAATGTTGGGAGC 277028
 QY 301 CCGGATGATTTGCGTTGTTGTTCTTGTGTTATTAATACGCTACGAATGTTGGGAGC 360
 Db 277029 CCGGATGATTTGCGTTGTTGTTCTTGTGTTATTAATACGCTACGAATGTTGGGAGC 277088
 QY 361 TGTGGGCGAGGCGCTTGTGGGCGCATGCGCTACCAATTTGCTGTAACCTTCCGCA 420
 Db 277089 TGTGGGCGAGGCGCTTGTGGGCGCATGCGCTACCAATTTGCTGTAACCTTCCGCA 277148
 QY 421 CGGATTCATAGGTTTTCATTCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 277149 CGGATTCATAGGTTTTCATTCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 277208
 QY 481 TTGCGTTTTCGCGCTTATGAGCAAGGCTTACGAGCTACGAGGCGAGGCGAGGCTG 540
 Db 277209 TTGCGTTTTCGCGCTTATGAGCAAGGCTTACGAGCTACGAGGCGAGGCGAGGCTG 277268
 QY 541 TTGCGGCTGCGGCACTTATGCAATGCAATGCGTTTACCTTCTGAGGATTTCCGCGG 600
 Db 277269 TTGCGGCTGCGGCACTTATGCAATGCAATGCGTTTACCTTCTGAGGATTTCCGCGG 277328
 QY 601 TCTCACCGAGTGGCGATGCGAGAGTTGATTTCTTTGAAATTAAGGCGCTCGAGTTGC 660
 Db 277329 TCTCACCGAGTGGCGATGCGAGAGTTGATTTCTTTGAAATTAAGGCGCTCGAGTTGC 277388
 QY 661 CTTTGTCTCTCTTCTTGTGAGCTGATCTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 720
 Db 277389 CTTTGTCTCTCTTCTTGTGAGCTGATCTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 277448
 QY 721 TTCTGTCTCTCTCTGAGGTTTGTGAGCTTCAACATTTCTTGTGATTTCAAGGCTGAGC 780
 Db 277449 TTCTGTCTCTCTCTGAGGTTTGTGAGCTTCAACATTTCTTGTGATTTCAAGGCTGAGC 277508
 QY 781 CCGATTTGCGGCGCTGATCTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 840
 Db 277509 CCGATTTGCGGCGCTGATCTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 277568
 QY 841 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 900
 Db 277569 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 277628
 QY 901 GTGATTAATTTTCCGCTCCGCGGCGTTCCGTTCTTAATCTTAACCCCTACGTGAATCA 960
 Db 277629 GTGATTAATTTTCCGCTCCGCGGCGTTCCGTTCTTAATCTTAACCCCTACGTGAATCA 277688
 QY 961 CAATTGTGGGCAAAATGCGATGTGATGCGACGAGAAATCTTGCAATTTTGAACGCA 1020
 Db 277689 CAATTGTGGGCAAAATGCGATGTGATGCGACGAGAAATCTTGCAATTTTGAACGCA 277748
 QY 1021 TCAAGCTTTGCGAGCAATGAGATGATGAGAACTCTTAACCTTTGTTCTATGCGCTT 1080
 Db 277749 TCAAGCTTTGCGAGCAATGAGATGATGAGAACTCTTAACCTTTGTTCTATGCGCTT 277808
 QY 1081 GCGATTACAGTGTGCGGCACTCTTCTGCGGTCGACGCACTTGTGAGCGTTGCGGCT 1140
 Db 277809 GCGATTACAGTGTGCGGCACTCTTCTGCGGTCGACGCACTTGTGAGCGTTGCGGCT 277868
 QY 1141 GGCACCATGTTTTTGTGGAATGATCTTTTCTAAACTGATTAATTAACAAAT 1200
 Db 277869 GGCACCATGTTTTTGTGGAATGATCTTTTCTAAACTGATTAATTAACAAAT 277928
 QY 1201 CCGCATGCTCTCAATTTGAAGGGATGCGAATTTTGAAGAACTGAAGAAAGCTTAAG 1260
 Db 277929 CCGCATGCTCTCAATTTGAAGGGATGCGAATTTTGAAGAACTGAAGAAAGCTTAAG 277988

```

gene
CDS
5354..7489
/gene="Cg10006"
5354..7489
/gene="Cg10006"
/EC_number="5.99.1.3"
/Note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:Toprim domain
PF02518:Histidine kinase-, DNA gyrase B-, phytochrome-like
ATPase
TIGR01055:parE_Gneg: DNA topoisomerase IV, B subunit
TIGR01058:parE_gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/translation="transl_table=11"
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="NRGTTMGPKYRMRKLYRISSECSLAVANTHENYDASITILE
GLEAVRKPGMTIGSTGPRGLHLIWEVDNSVDAMAGHATKVEVTLLEPGQVYD
DORGI.PVMEHPSGAPVQVWMTQLHAGKFDSDSYASGGLHGVISVNLSTREVA
DILKHKWYQNFERSVDELEEGNAGGTGTTIRFMDAEI FETTEDPFTISRRLQ
EKAFNLKGLITLTDNRATDELEALAEQGETEIDEINTELEVEITDAPK
KKRKKKKI FHYPNGLDYVYLRNKTNIHPSIVSEALGDHDEVANQMSYK
BSVHPFANTIRREGTEHEGRSALTSLMRVPSHEHLLEKKNLIGDRCGLSA
VTSVVDPPFGQRTTKIKNTETIS FVQRANBEIGWILANPRAKATNRYGSA
SNFQALPLRGRKILVERKARDLGLPKLADCKSKDPESELTLEGGDAGSASAGSD
QNRQALPLRGRKILVERKARDLGLPKLADCKSKDPESELTLEGGDAGSASAGSD
ADADYDGHIAATLTLFLFRFMDLVARHGVLAQPLPKWKQGEFGVSEBERD
EOLNGLAGRKINDGDIQRYKGLGEMNASLMTETMDPTVRLIRVDITDAQRADE
LFSILMGDDVVARSPITRNADVRFID"
7830..8798
/gene="Cg10007"
7830..8798
/gene="Cg10007"
/codon_start=1
/translation="transl_table=11"
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="KIQITILKIGIKAKQRMKATLITDILGHVSTPTTHAVP
LDPGSSTPVQIWAADKDSQVWMLWPGFMGYYRPLAALNKGHVAIGELRGQG
OSAKAKRSQWGYHDLASVDFPQIAAKKALDIEGHMFLSHSGQGISCLFAA
REBAKYNLRALFGVAGSPFRPTSPKNGKLGAVLGGIGIIGHIVGFWGKVLGK
DLVYGROSGTHMRWRFRHKNLSLDLTAODINVEVWKVSIPIITSRCDDDCP
QASIDGLAFVPAQIKMTEIPEALGHNRMAREPSTVYLFLEQAL"
complement(8795..9466)
/gene="Cg10008"
complement(8795..9466)
/gene="Cg10008"
/codon_start=1
/translation="transl_table=11"
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translation="MSNENKRLDPAQKSLPTRNAMPVGFIFMSFAGITVHAGLG
TTTISLPLVNTAGSLSGWTTIFNGSMILCOITLVLSKPTPLVQILALFJGF
LQDLSLQITWAGTDNYPVAMIMVIVSTILMSIGITVLPNITPLAAGISALVKG
FVVERFGTMKQIVDMTFVVAAILSWITMGILIVREGVFAAFIFGFEVROWRLYL
RSIGH"
complement(9471..9914)
/gene="Cg10009"
complement(9471..9914)
/gene="Cg10009"
/codon_start=1
/translation="transl_table=11"
/product="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"
/translation="MTSLRDLVSTNDAALADMSAFIEETVKGSGISGIALKQMG

```

```

AAATKIDSDIVAKSRRLPETADSLDGLMOEYONGTAPADFGALFANSSTALDALS
VADKIDABESINBGLKRYKGYGRKAAKXVLEQSLPAGLIEKNAQ"
gene
CDS
complement(10104..11177)
/gene="Cg10010"
complement(10104..11177)
/gene="Cg10010"
/codon_start=1
/translation="transl_table=11"
/product="Hypothetical protein"
/protein_id="BAB97403.1"
/db_xref="GI:21322774"
/translation="MEMLAQRILPHITAPHLPVAVRPHQIRATDADRTSLQRDLELS

Query Match      100.0% Score 1271; DB 1; Length 34000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGCGATCAATGGAATCTAGCTTCATATATGCACAATAGCTTAGTGAGTCCGCAAC 60
Db      276729 GCGCGATCAATGGAATCTAGCTTCATATATGCACAATAGCTTAGTGAGTCCGCAAC 276788
QY      61 TGGCAACAAACATACCGGCAATGTGTGATGATGATGTCACAAAACGCAAGAGAT 120
Db      276789 TGGCAACAAACATACCGGCAATGTGTGATGATGATGTCACAAAACGCAAGAGAT 276848
QY      121 TCAATCAAGCCTGAGAGGTGTGGCAATCAAGGAGCCTTGAAACGATGATTAAGGTTA 180
Db      276849 TCAATCAAGCCTGAGAGGTGTGGCAATCAAGGAGCCTTGAAACGATGATTAAGGTTA 276908
QY      181 TCGCGCTAGGAATGCGCGAAGGTCTAAACCTCCCTTGCTGCGAGGTTGGGCAATGA 240
Db      276909 TCGCGCTAGGAATGCGCGAAGGTCTAAACCTCCCTTGCTGCGAGGTTGGGCAATGA 276968
QY      241 CCGGATGTATTCGTTTGGTCTCTGTGTTATTAACGAGCTACGAATGTGGGCAAGC 300
Db      276969 CCGGATGTATTCGTTTGGTCTCTGTGTTATTAACGAGCTACGAATGTGGGCAAGC 277028
QY      301 CCCACTGTTTTCGGGCTGATTTTTCGGGCTCCACGAAATGCTGTCATGCGCTCGT 360
Db      277029 CCCACTGTTTTCGGGCTGATTTTTCGGGCTCCACGAAATGCTGTCATGCGCTCGT 277088
QY      361 TGTGGGCGACGCGCCCTGCGGCGCAATCGCGCTCACGCAATTCGAGTGAATCTCCGCA 420
Db      277089 TGTGGGCGACGCGCCCTGCGGCGCAATCGCGCTCACGCAATTCGAGTGAATCTCCGCA 277148
QY      421 CGTATTCATGCGTTTTCATTTCCCGCTGCAATGTGTGTCAAAACCCCATTCGCCGTTCTA 480
Db      277149 CGTATTCATGCGTTTTCATTTCCCGCTGCAATGTGTGTCAAAACCCCATTCGCCGTTCTA 277208
QY      481 TTGGTTCGCGCTTATCGACGAGGCTACGAGTCACTGCGGCGACGCGCAGGCTG 540
Db      277209 TTGGTTCGCGCTTATCGACGAGGCTACGAGTCACTGCGGCGACGCGCAGGCTG 277268
QY      541 GTGGGCTGCGGACTTATCTCAATGCAATAGGTTTCACTCTACTGCGTATTCGGCGG 600
Db      277269 GTGGGCTGCGGACTTATCTCAATGCAATAGGTTTCACTCTACTGCGTATTCGGCGG 277328
QY      601 TCTCACCAGGAGTGGCAATTCGAGAGTGAATCTCTTGAATTAAGGGCTCGAGTTGCG 660
Db      277329 TCTCACCAGGAGTGGCAATTCGAGAGTGAATCTCTTGAATTAAGGGCTCGAGTTGCG 277388
QY      661 CTTTGGCTCTCTCTTTGTACAGCTGACTTTGGATTCCTGCGGAAACGAAAAGCAGATCCC 720
Db      277389 CTTTGGCTCTCTCTTTGTACAGCTGACTTTGGATTCCTGCGGAAACGAAAAGCAGATCCC 277448
QY      721 TTTCTGTGCTGCGCAGGTTTGAAGCTTCAACATGCTCTTGGTAAATTCAGGTCAGGC 780
Db      277449 TTTCTGTGCTGCGCAGGTTTGAAGCTTCAACATGCTCTTGGTAAATTCAGGTCAGGC 277508
QY      781 CCTATTGGCGGCGGTGTGATCTTCTTGGGTCGTGTGACCAATCGGTAATCTTCTTGGG 840
Db      277509 CCTATTGGCGGCGGTGTGATCTTCTTGGGTCGTGTGACCAATCGGTAATCTTCTTGGG 277568

```


QY 901 GTCATTACTTTGGCGCTCCGGGCGGTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
 DB 901 GTCATTACTTTGGCGCTCCGGGCGGTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
 QY 961 CAATTTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTGAACGCA 1020
 DB 961 CAATTTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTGAACGCA 1020
 QY 1021 TCAAGTTTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTGAACGCA 1080
 DB 1021 TCAAGTTTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTGAACGCA 1080
 QY 1081 GCGATTAGATGATGCGCGATCTTCTTGCGCGGTGACGCACTTGTTGAGCGTTGCGGCT 1140
 DB 1081 GCGATTAGATGATGCGCGATCTTCTTGCGCGGTGACGCACTTGTTGAGCGTTGCGGCT 1140
 QY 1141 GCGACCATGCTTTTGTGAGTGTGATCTTTTCTTAACTGATTAATTAACAAAT 1200
 DB 1141 GCGACCATGCTTTTGTGAGTGTGATCTTTTCTTAACTGATTAATTAACAAAT 1200
 QY 1201 CCGCATGCGCTCAATTTGAAGGAGATGGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 DB 1201 CCGCATGCGCTCAATTTGAAGGAGATGGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 7
 LOCUS AP005274 340000 bp DNA linear ECT 08-AUG-2002
 DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
 ACCESSION AP005274 BAA000036
 VERSION AP005274.1 GI:21322764
 KEYWORDS
 ORGANISM
 Corynebacterium glutamicum ATCC 13032
 Corynebacterium glutamicum ATCC 13032
 Bacteria; Actinobacteriae; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1 Nakagawa, S.
 Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
 TITLE
 Unpublished
 JOURNAL
 2 (bases 1 to 340000)
 AUTHORS
 Nakagawa, S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Aabhi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@kaken.gen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
 COMMENT
 This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. and Kitasato University.
 FEATURES
 source
 1. 340000
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 1. 1575
 /gene="Cg10001"
 1. 1575
 /gene="Cg10001"
 /note="PB00308: Bacterial dnaA protein
 TIGR00362: DnaA: chromosomal replication initiator protein
 DnaA"
 /codon_start=1
 /transl_table=1
 /product="ATPase involved in DNA replication initiation"
 /protein_id="BAB97394.1"
 /db_xref="GI:21322765"

gene
 CDS
 3585..4769
 /gene="Cg10004"
 3585..4769
 /gene="Cg10004"
 /note="TIGR00611: recf: RecF protein"
 /codon_start=1
 /transl_table=1
 /product="Recombinational DNA repair ATPase"
 /protein_id="BAB97397.1"
 /db_xref="GI:21322768"
 /transl_table=1
 /note="WHIRLS:LELDRYSWDELKYDELPGITVFIRNGFGKTNIVEAIG
 YLAHSSHRVSDAPLRAHAENARVAVAVNGRELAHLILKPAANAQASINRKY
 RTRELLGVKTVLPAPEDALVKGPAERARVLDLIIARCPRMAGVADYKVKQ
 RNALIKTATIALRGVGTGEGAALSTLDWDGLRAGVVAAPFALLNEGPITY
 EATTTAPBSRPAANVYKTTIDGLISQEPFPGIIEATLITELAKROREIRGSSL
 VGHRRDVIDLGDQPAKGPASGRTWFRSLRIAPFNILKSDGDPDIIIDYVSE
 LDAGRRKLVGIAQVEVQVLTAAVDDLPENILKVLTAQHTVTDGTGRLSLDDV
 QP"
 4766..5302
 /gene="Cg10005"
 4766..5302
 /gene="Cg10005"
 /codon_start=1
 /transl_table=1
 /product="Hypothetical protein"
 /protein_id="BAB97398.1"
 /db_xref="GI:21322769"
 /transl_table=1
 /note="MTDPIQAFERIRABAMRRANGVPLDKNDAPRRPPAPKGVEX

/transl_table=1
 /product="Hypothetical protein"
 /protein_id="BAB97395.1"
 /db_xref="GI:21322766"
 /transl_table=1
 /note="PF00712: DNA polymerase III beta subunit, N-terminal domain
 PF02767: DNA polymerase III beta subunit, C-terminal domain
 PF02768: DNA polymerase III beta subunit, C-terminal domain
 TIGR00663: dnan: DNA polymerase III, beta subunit"
 /codon_start=1
 /transl_table=1
 /product="DNA polymerase III beta subunit"
 /protein_id="BAB97396.1"
 /db_xref="GI:21322767"
 /transl_table=1
 /note="MESONS:PFVAVRVEDIIVTAVAVARNLPTKQGPVLRAMLITAD
 EGELAGDYDSTRVRLSABVSQPRIVAVAKLSEITGSLPNKRPDRIGSKAFV
 TCGSSRPELPLLDYPMPLKIPAGTGINAKLFEVAVSQAAGKXDSPLMTLTV
 SMEIVQNIPLATDPRFRLALRFEWANDPELVNLIIPARTLDNASSLDSGLND
 IDIAGVTGQVGRGELFGVHTDNRRTLRLLADPENIPLPKERTALASVEIAYL
 DAIRVSLVARNAGIVLHFSBGVILITAGATGAGAEETLPCAFPGKELTIAFNQY
 LKDGSLVPTSRVAFGTSPRAIMIPPEMPSANENGIFQTPPTVFTYLLMPVRL
 PG"
 3585..4769
 /gene="Cg10004"
 3585..4769
 /gene="Cg10004"
 /note="TIGR00611: recf: RecF protein"
 /codon_start=1
 /transl_table=1
 /product="Recombinational DNA repair ATPase"
 /protein_id="BAB97397.1"
 /db_xref="GI:21322768"
 /transl_table=1
 /note="WHIRLS:LELDRYSWDELKYDELPGITVFIRNGFGKTNIVEAIG
 YLAHSSHRVSDAPLRAHAENARVAVAVNGRELAHLILKPAANAQASINRKY
 RTRELLGVKTVLPAPEDALVKGPAERARVLDLIIARCPRMAGVADYKVKQ
 RNALIKTATIALRGVGTGEGAALSTLDWDGLRAGVVAAPFALLNEGPITY
 EATTTAPBSRPAANVYKTTIDGLISQEPFPGIIEATLITELAKROREIRGSSL
 VGHRRDVIDLGDQPAKGPASGRTWFRSLRIAPFNILKSDGDPDIIIDYVSE
 LDAGRRKLVGIAQVEVQVLTAAVDDLPENILKVLTAQHTVTDGTGRLSLDDV
 QP"
 4766..5302
 /gene="Cg10005"
 4766..5302
 /gene="Cg10005"
 /codon_start=1
 /transl_table=1
 /product="Hypothetical protein"
 /protein_id="BAB97398.1"
 /db_xref="GI:21322769"
 /transl_table=1
 /note="MTDPIQAFERIRABAMRRANGVPLDKNDAPRRPPAPKGVEX

QY 721 TTCTGTGCTGCGAGGTTGAGCTTCAACATGCTCTTGTGTAATTCAGGTACAGC 780
 DB 721 TTCTGTGCTGCTGCGAGGTTGAGCTTCAACATGCTCTTGTGTAATTCAGGTACAGC 780
 QY 781 CCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACCACTCCGGTAATCTTCTTGGG 840
 DB 781 CCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACCACTCCGGTAATCTTCTTGGG 840
 QY 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGCA 900
 DB 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGCA 900
 QY 901 GTGATTACTTTTGGCGCTCCGCGCGGCTCCGTTCTTAATCTCTTAAGCCCTTAAGTAATCA 960
 DB 901 GTGATTACTTTTGGCGCTCCGCGCGGCTCCGTTCTTAATCTCTTAAGCCCTTAAGTAATCA 960
 QY 961 CAATTGTGGGGAATAATGGCGATGCGATGCGAGGAGATCTTGCCATTTGACCGCA 1020
 DB 961 CAATTGTGGGGAATAATGGCGATGCGATGCGAGGAGATCTTGCCATTTGACCGCA 1020
 QY 1021 TCAAGCTTTGCGAGCAATGCGATGATCTGAAGACTTAACCTTTGCTGATTCGCGCT 1080
 DB 1021 TCAAGCTTTGCGAGCAATGCGATGATCTGAAGACTTAACCTTTGCTGATTCGCGCT 1080
 QY 1081 GCGATTACAGTGTGGCGGATCTTCTTGGCGGTCAGACGACCTTTGTAGCGCTGGCGCT 1140
 DB 1081 GCGATTACAGTGTGGCGGATCTTCTTGGCGGTCAGACGACCTTTGTAGCGCTGGCGCT 1140
 QY 1141 GGCACCATGCTTTTGTGTGACTGTGATCTTTTCTAAACTGTGCTAAATTAACAAAT 1200
 DB 1141 GGCACCATGCTTTTGTGTGACTGTGATCTTTTCTAAACTGTGCTAAATTAACAAAT 1200
 QY 1201 CCGCATGCGCTCAATTGAAAGGAGATGCGGATTTTCTTAAGGAACCTTAAGAAAGGCTTAG 1260
 DB 1201 CCGCATGCGCTCAATTGAAAGGAGATGCGGATTTTCTTAAGGAACCTTAAGAAAGGCTTAG 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 6
 BD014994 1271 bp DNA linear PAT 27-AUG-2002
 LOCUS Nucleotide sequence encoding the discharge of branched amino acid,
 DEFINITION method of isolating the same and utilization thereof.
 ACCESSION BD014994
 VERSION BD014994.1 GI:22555801
 KEYWORDS JP 200116978-A/4.
 SOURCE unidentified
 ORGANISM unclassified.
 1 (bases 1 to 1271)
 Kernerukunehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 200116978-A 4 26-JUN-2001;
 DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC13032
 COMMENT
 JOURNAL
 AUTHORS
 TITLE
 JOURNAL
 DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC13032
 PN JP 200116978-A/4
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KERNERKUNEHITO, HERMANN SAHM, LOTMAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08, C12N15/09, C12R1/15, (C12N1/21, C12R1/15), (C12P13/06, C12R1/15),
 PC C12N15/00, C12R1/15)
 PC (C12N15/00, C12R1/15)
 CC bnf
 CC bnf
 FH key

Location/Qualifiers

FEATURES
 source
 1.1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATGCAATGGAATCTAGCTTCATATATTCACAAATAGCTAGTGGTCCCAAC 60
 DB 1 GCGGATGCAATGGAATCTAGCTTCATATATTCACAAATAGCTAGTGGTCCCAAC 60
 QY 61 TGGCAACAAACTACCCGCAATTGTGTATGATTTAGTGTGCAAAAACCGAAGAT 120
 DB 61 TGGCAACAAACTACCCGCAATTGTGTATGATTTAGTGTGCAAAAACCGAAGAT 120
 QY 121 TCAATGAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACGATATAAGGTTA 180
 DB 121 TCAATGAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACGATATAAGGTTA 180
 QY 181 TCGGCGCTACGAATGCGCGCAAGTCTAAAACTCCCTTGCTGCAAGTTGGGATGTA 240
 DB 181 TCGGCGCTACGAATGCGCGCAAGTCTAAAACTCCCTTGCTGCAAGTTGGGATGTA 240
 QY 241 CCGGATGATTTGCTGCTTGTGTCTCTTGTATTAATACGGCTAGAAATGCTGGGACG 300
 DB 241 CCGGATGATTTGCTGCTTGTGTCTCTTGTATTAATACGGCTAGAAATGCTGGGACG 300
 QY 301 CCGATTTGATTTGCTGCTTGTGTCTCTTGTATTAATACGGCTAGAAATGCTGGGACG 360
 DB 301 CCGATTTGATTTGCTGCTTGTGTCTCTTGTATTAATACGGCTAGAAATGCTGGGACG 360
 QY 361 TGTGGGCGAGCGCCCTGCGGCGCATGCGGCTCACCAATGCTGTGATCTTCGCCA 420
 DB 361 TGTGGGCGAGCGCCCTGCGGCGCATGCGGCTCACCAATGCTGTGATCTTCGCCA 420
 QY 421 CGATTTCTAGCGTTTCAATTCGCGCTGATGTGTCAAAAACCGATGCGCGTTCTA 480
 DB 421 CGATTTCTAGCGTTTCAATTCGCGCTGATGTGTCAAAAACCGATGCGCGTTCTA 480
 QY 481 TTGGTTTGGCGCTTATGACAGAGCTTACGAGTCACTGCGGCGCAGCGCGAGCTG 540
 DB 481 TTGGTTTGGCGCTTATGACAGAGCTTACGAGTCACTGCGGCGCAGCGCGAGCTG 540
 QY 541 GTGGGCGTGGGCACTTATCTCAATAGCAATAGAGCTTCACTGCTAGTGGTATTCGCGG 600
 DB 541 GTGGGCGTGGGCACTTATCTCAATAGCAATAGAGCTTCACTGCTAGTGGTATTCGCGG 600
 QY 601 TCTACCGAGTGGCGATGCGAGAGTGAATCTTTTGAATTAAGGCGCTCGAGTTGC 660
 DB 601 TCTACCGAGTGGCGATGCGAGAGTGAATCTTTTGAATTAAGGCGCTCGAGTTGC 660
 QY 661 CTTTGGCTCTCTTTTTCACGCGTGAATTTGATTTCTGCGGAGCAAAAAGAGATCCC 720
 DB 661 CTTTGGCTCTCTTTTTCACGCGTGAATTTGATTTCTGCGGAGCAAAAAGAGATCCC 720
 QY 721 TTCTGTGCTGCTGAGGTTGAGCTTCAACATGCTCTTGTGTAATTCAGGTACAGC 780
 DB 721 TTCTGTGCTGCTGAGGTTGAGCTTCAACATGCTCTTGTGTAATTCAGGTACAGC 780
 QY 781 CCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACCACTCCGGTAATCTTCTTGGG 840
 DB 781 CCTATTGCGGCGCTGCTGATCTTCTTGGGTCTGTGACCACTCCGGTAATCTTCTTGGG 840
 QY 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGCA 900
 DB 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGCA 900

QY 421 CGTATTCATGCGTTTTCATTCGCCGCTGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 DB 421 CGTATTCATGCGTTTTCATTCGCCGCTGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTCGGTTTCGGGCTTATGAGAGAACCCGACATCTGGGGGCGAGGCGGCTG 540
 DB 481 TTCGGTTTCGGGCTTATGAGAGAACCCGACATCTGGGGGCGAGGCGGCTG 540
 QY 541 GTCGGGCTGCGCATTCATTCATGCAATAGCAGTTTCACTCTCACTGGGATTTGGCGG 600
 DB 541 GTCGGGCTGCGCATTCATTCATGCAATAGCAGTTTCACTCTCACTGGGATTTGGCGG 600
 QY 601 TCTCACCGAGATGCGCATGCGAGATTCCTTTTGAATAAGGCGCTCGAATGCG 660
 DB 601 TCTCACCGAGATGCGCATGCGAGATTCCTTTTGAATAAGGCGCTCGAATGCG 660
 QY 661 CCTTTCCTCTCTCTTGTGACGCTGACTTGGATTCCTGCGAGCAAGAAAGCAGATCCC 720
 DB 661 CCTTTCCTCTCTCTTGTGACGCTGACTTGGATTCCTGCGAGCAAGAAAGCAGATCCC 720
 QY 721 TTTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 DB 721 TTTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 QY 781 CCTATTTGGGGGCGCTGATCTTCTGGGTCGTTGACCATCCGATCTTCTTCTGGG 840
 DB 781 CCTATTTGGGGGCGCTGATCTTCTGGGTCGTTGACCATCCGATCTTCTTCTGGG 840
 QY 841 AAGAGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTATGCA 900
 DB 841 AAGAGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTATGCA 900
 QY 901 GTCATTAATTTGGGCTGCGGCGGCTCGGCTTCTTAATCCCTTAAGCCCTAGCATGCA 960
 DB 901 GTCATTAATTTGGGCTGCGGCGGCTCGGCTTCTTAATCCCTTAAGCCCTAGCATGCA 960
 QY 961 CAATTTGGGGGCAAAATGGCGATGTCATGTCAGCAGAGATCTTCCATTTTGAACCGCA 1020
 DB 961 CAATTTGGGGGCAAAATGGCGATGTCATGTCAGCAGAGATCTTCCATTTTGAACCGCA 1020
 QY 1021 TCAAGCTTTGCGAGCATGTCATGTCAGCAGAGATCTTCCATTTTGAACCGCT 1080
 DB 1021 TCAAGCTTTGCGAGCATGTCATGTCAGCAGAGATCTTCCATTTTGAACCGCT 1080
 QY 1081 GCGATTAAGTGTGCGGCTGCTTCTTGGCGGTCAGCAGCATCTTGTGAGCGCTGCGCT 1140
 DB 1081 GCGATTAAGTGTGCGGCTGCTTCTTGGCGGTCAGCAGCATCTTGTGAGCGCTGCGCT 1140
 QY 1141 GGCACCATGCTTTTGTGAGTGTGATCTTTTCTAAATGCGCTAAATTAACAAAT 1200
 DB 1141 GGCACCATGCTTTTGTGAGTGTGATCTTTTCTAAATGCGCTAAATTAACAAAT 1200
 QY 1201 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 DB 1201 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 4
 AX137714 1271 bp DNA linear PART 30-MAY-2001
 LOCUS AX137714
 DEFINITION Sequence 6 from Patent EP1096010.
 ACCESSION AX137714
 VERSION AX137714.1 GI:14273893
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

AUTHORS Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 TITLE Nucleotide sequences coding for branched-chain amino acids export
 JOURNAL proteins, method for isolating them and their use
 Patent: EP 1096010-A 6 02-MAY-2001;
 Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 location/Qualifiers
 1..1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC13032"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39927.1"
 /db_xref="GI:14273894"
 /db_xref="REMBL:CAC39927"
 /translation="MKTQBIHSLVSPSKALBPDDKGRRYEINQGLKSLAAGL
 GMPDIAFGLVIOYGEWMAFLPGLIPAGSTEMLVIALVGAAPLGAIALTL
 VNRHVPVAFSPFLHYKPIARFVSPALIDBAVTAAPAGVGAAPLISQIAFH
 SYVFGSTVQAIAELIPEIKGERLALSLPVTLLDSCRKQIPSLILGLSPFI
 ALVITPQALFPAALIFLGILTRIFFLGKAK"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39928.1"
 /db_xref="GI:14273895"
 /db_xref="REMBL:CAC39928"
 /translation="MTTDFSCILLVAVCAVITPALRANPFLIKLPRESQFVKKAM
 WMPAGIATITASTFRSNAIDKLTFLGIAVAITVVAHLGSRLLSVAGTIVFV
 GLVNLFF"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATCAATGGAATGATCTTCAATATTTGACATATAGCCTTATGAGTGGCGCAAC 60
 DB 1 GCGGATCAATGGAATGATCTTCAATATTTGACATATAGCCTTATGAGTGGCGCAAC 60
 QY 61 TGGCAACAAACTACCCGCGAATGTGTGATGTATGTATGTGCAAAAAACGCAAGAT 120
 DB 61 TGGCAACAAACTACCCGCGAATGTGTGATGTATGTATGTGCAAAAAACGCAAGAT 120
 QY 121 TCATTCAAGCTGGAAGTGTGCGCATCCAGAGCGCTGGAACCAAGATATAAGGTTA 180
 DB 121 TCATTCAAGCTGGAAGTGTGCGCATCCAGAGCGCTGGAACCAAGATATAAGGTTA 180
 QY 181 TGGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCCTGCTGCGAGTTTGGCATGTA 240
 DB 181 TGGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCCTGCTGCGAGTTTGGCATGTA 240
 QY 241 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 300
 DB 241 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 300
 QY 241 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 300
 DB 241 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 300
 QY 301 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 360
 DB 301 CCGCATTTGATGTCGTTGCTCTTGTATTAATCAATGCGCTACGATGTCGCGAGC 360
 QY 361 TGTGGGCGGAGCGCCCTGCGGCGCATGCGGCTACACCAATGTGTGGAATCTCCGCA 420
 DB 361 TGTGGGCGGAGCGCCCTGCGGCGCATGCGGCTACACCAATGTGTGGAATCTCCGCA 420
 QY 421 CGTATTCATGCGTTTTCATTCGCGCTGATGTCGTCGCAAAACCCCATTTGCCGTTCTA 480
 DB 421 CGTATTCATGCGTTTTCATTCGCGCTGATGTCGTCGCAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTCGGTTTCGGGCTTATGAGAGAACCCGACATCTGGGGGCGAGGCGGCTG 540

Db 301 CCCACTGTTTTCGGGCGTGAATTTTCGGGCGGCTCAACCGAAATGCTGATCGCCCTCGT 360
 QY 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGCCGCTCACACATTCGCTGGAATCTTCGCGCA 420
 Db 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGCCGCTCACACATTCGCTGGAATCTTCGCGCA 420
 QY 421 CGTATTCATGAGGTTTTCATTCGCGCGATGCTGCTCAAAAACCCCATTCGCCGTTCTA 480
 Db 421 CGTATTCATGAGGTTTTCATTCGCGCGATGCTGCTCAAAAACCCCATTCGCCGTTCTA 480
 QY 481 TTGCGTTTTCGCGCTTATGAGAGAACTTACGCACTGCTGCGGCGCAAGCCGCGAGCTG 540
 Db 481 TTGCGTTTTCGCGCTTATGAGAGAACTTACGCACTGCTGCGGCGCAAGCCGCGAGCTG 540
 QY 541 GTCGCGGTCGAGCTTATGATGCAATGCAATGCTTCACTGCTGCTGCTGCTGCTGCTGCTG 600
 Db 541 GTCGCGGTCGAGCTTATGATGCAATGCAATGCTTCACTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 TCTCAACCGAGTGGCGATGCGAGAGTGAATTCCTTTGAATTAAGGCGCTCGAGTTGCG 660
 Db 601 TCTCAACCGAGTGGCGATGCGAGAGTGAATTCCTTTGAATTAAGGCGCTCGAGTTGCG 660
 QY 661 CCTTTGCTCTCTTTTGTGACGCTGACCTTTGATTCCTGCGCAAGAAAGAGAGATGCC 720
 Db 661 CCTTTGCTCTCTTTTGTGACGCTGACCTTTGATTCCTGCGCAAGAAAGAGAGATGCC 720
 QY 721 TTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTTGTGTATTCAGTCAAGTCAAGC 780
 Db 721 TTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTTGTGTATTCAGTCAAGTCAAGC 780
 QY 781 CCTATTGGGCGGCTGCTGATCTTCTTGGGCTGCTGTAACATCCGCTACTTCTTCTGGG 840
 Db 781 CCTATTGGGCGGCTGCTGATCTTCTTGGGCTGCTGTAACATCCGCTACTTCTTCTGGG 840
 QY 841 AAAGGCTGCTAAATGACATGATATTCCTGCTGATTCCTTGTGTGCTGCTGCTGCTGCTG 900
 Db 841 AAAGGCTGCTAAATGACATGATATTCCTGCTGATTCCTTGTGTGCTGCTGCTGCTGCTG 900
 QY 901 GTCATTTACTTTTGGGCTCCGCGGCGGTTCCGCTTCTTAATCTTAAGCCCTTAAGTGAATCA 960
 Db 901 GTCATTTACTTTTGGGCTCCGCGGCGGTTCCGCTTCTTAATCTTAAGCCCTTAAGTGAATCA 960
 QY 961 CAATTTGCGGCAAAATGCGATGCGATGCGAGAGAACTCTTGCAATTTGACCGCA 1020
 Db 961 CAATTTGCGGCAAAATGCGATGCGATGCGAGAGAACTCTTGCAATTTGACCGCA 1020
 QY 1021 TCAACGTTTTCGAGCAATGCGATGATCTGAAGACTTAACTTTGCTCATTCGCGTT 1080
 Db 1021 TCAACGTTTTCGAGCAATGCGATGATCTGAAGACTTAACTTTGCTCATTCGCGTT 1080
 QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGACGCACTTGTGAGCGCTGAGCCT 1140
 Db 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGACGCACTTGTGAGCGCTGAGCCT 1140
 QY 1141 GGCACCATGTTTGTGAGACTGATGATCTTTTCTAAACCTGATTAATTAACAAAT 1200
 Db 1141 GGCACCATGTTTGTGAGACTGATGATCTTTTCTAAACCTGATTAATTAACAAAT 1200
 QY 1201 CCGCATGCGCTCAATTTGAAGGGAATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 Db 1201 CCGCATGCGCTCAATTTGAAGGGAATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

RESULT 3
 AX137709 1271 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137709
 DEFINITION Sequence 1 from Patent EP1096010.
 ACCESSION AX137709
 VERSION AX137709.1 GI:14273886

KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 Location/Qualifiers
 1. 1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 101. 856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39923.1"
 /db_xref="EMBL:CAC39923"
 /translation="MQRDRIHSSLSLESPSKALBPDQKRYRETAQGLKTSIAAGL
 GMPTPIGAFGLVIVQYGEWMAPIFSGIPIAGSTMLVIALVGAAPGIALTLTL
 VNFRAHYAFSPFLHVKNPIDAFYFVPLIDBAVYAVTAPAGWAMBLISQIAFH
 SYVFGGLNVAIAELIPEIKGLERALSLEVTLLDISCRTKQIPSLIAGLSFTI
 ALVIVGQALFALLIFLGLITRYFPLGKAAK"
 853. 1179
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=11
 /protein_id="GI:14273888"
 /db_xref="EMBL:CAC39924"
 /translation="MTTPSCILIVAVCAVITFALRAVPLILKRLRSOPFGKAM
 WMPAGILAITASTFRSMAIDKTLIFGLIAVAITVVAHLGGRRLLSVGACTIVFV
 GLVNLV"

CDs
 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
 Query Match
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATCAATGGAATTAAGCTTCAATATTAATGCAATTAAGCTTAAGTGAAGTGGCAAC 60
 Db 1 GCGGATCAATGGAATTAAGCTTCAATATTAATGCAATTAAGCTTAAGTGAAGTGGCAAC 60
 QY 61 TGGCAACAAACCTACCGGCAATTTGTGATGATGATGATGATGATGATGATGATGATGATG 120
 Db 61 TGGCAACAAACCTACCGGCAATTTGTGATGATGATGATGATGATGATGATGATGATGATG 120
 QY 121 TCATTAAGCTGAGGCTGTGCCATCAAGGAGCCCTGGAACCGAGATATTAAGTTA 180
 Db 121 TCATTAAGCTGAGGCTGTGCCATCAAGGAGCCCTGGAACCGAGATATTAAGTTA 180
 QY 181 TCGGCGCTAGGAATGCGGCAAGGCTTAATAAACCCTCCCTGTGCGAGGTTGGGCAATGA 240
 Db 181 TCGGCGCTAGGAATGCGGCAAGGCTTAATAAACCCTCCCTGTGCGAGGTTGGGCAATGA 240
 QY 241 CCGCATTTGATTTGCGTTGCTCTTGTGTTATTAATGATGATGATGATGATGATGATGATG 300
 Db 241 CCGCATTTGATTTGCGTTGCTCTTGTGTTATTAATGATGATGATGATGATGATGATGATG 300
 QY 301 CCGCATTTTTCGCGGCTGATTTTCGCGGCTCAACCGAAATGCTGCTCATTCGCCCTGCT 360
 Db 301 CCGCATTTTTCGCGGCTGATTTTCGCGGCTCAACCGAAATGCTGCTCATTCGCCCTGCT 360
 QY 361 TGTGGGCGGAGCGCCCTGGGCGCATCGGCTCAACCAATTTGCTGTAAGTTCGCGCA 420
 Db 361 TGTGGGCGGAGCGCCCTGGGCGCATCGGCTCAACCAATTTGCTGTAAGTTCGCGCA 420

FEATURES		Location/Qualifiers	
source		1..1271	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Query Match	100.0%;	Score 1271;	DB 6; Length 1271;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1271;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
1	GGCGGATCAATGGAATCTAGCTTCAATATATGCAACAATAGCTAGTGGTGGCGCAAC	60	
1	GGCGGATCAATGGAATCTAGCTTCAATATATGCAACAATAGCTAGTGGTGGCGCAAC	60	
61	TGGCAAAAACCTACCCGGCAATTTGTGTGATGATTTAGTGTGCAAAAAACGCAAGAT	120	
61	TGGCAAAAACCTACCCGGCAATTTGTGTGATGATTTAGTGTGCAAAAAACGCAAGAT	120	
121	TCATTCAAGCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA	180	
121	TCATTCAAGCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA	180	
181	TGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGA	240	
181	TGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGA	240	
241	CCGATGATGATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGATGGGCAAC	300	
241	CCGATGATGATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGATGGGCAAC	300	
301	CCCACTGTTTCCGGCTGATTTTCGGCGGCTCCACCGAATGCTGGTCACTCCCTCGT	360	
301	CCCACTGTTTCCGGCTGATTTTCGGCGGCTCCACCGAATGCTGGTCACTCCCTCGT	360	
361	TGTTGGGCGAGCGCCCTGGCGGCGCATCGGCTCAACACATTCCTGGTGAACCTTCGGCA	420	
361	TGTTGGGCGAGCGCCCTGGCGGCGCATCGGCTCAACACATTCCTGGTGAACCTTCGGCA	420	
421	CGATATTCATGCGTTTCAATCCCGCTGCAATGTGTGCAAAAAACCCCATTCGCCGTTCTA	480	
421	CGATATTCATGCGTTTCAATCCCGCTGCAATGTGTGCAAAAAACCCCATTCGCCGTTCTA	480	
481	TTCCGTTTCCGCTTATCGAAGGCTACGAGTACGCGGCCAGCGCCGCAAGC	540	
481	TTCCGTTTCCGCTTATCGAAGGCTACGAGTACGCGGCCAGCGCCGCAAGC	540	
541	GTCGCGGTGCGCACTTATCTCAATGCAATAGCGTTTCACTCTCACTGGTATTCGCGG	600	
541	GTCGCGGTGCGCACTTATCTCAATGCAATAGCGTTTCACTCTCACTGGTATTCGCGG	600	
601	TCTCAACCGGAGTGGCGATCGCAAGTGAATTCCTTTGAAATTAAGGCCCTTCGAGTTCCG	660	
601	TCTCAACCGGAGTGGCGATCGCAAGTGAATTCCTTTGAAATTAAGGCCCTTCGAGTTCCG	660	
661	CCTTGCTCTCTCTTTGTCAGCTGACTTTGATTTCTGCGCAAGCAAAAAAGCAATCCC	720	
661	CCTTGCTCTCTCTTTGTCAGCTGACTTTGATTTCTGCGCAAGCAAAAAAGCAATCCC	720	
721	TTCTGTGCTGCTCGAGTTTGAAGTTCAATTCGCTTGTGTGTAATTCAGGTCAGGC	780	
721	TTCTGTGCTGCTCGAGTTTGAAGTTCAATTCGCTTGTGTGTAATTCAGGTCAGGC	780	
781	CCATTGTCGCGGCTGATCTTCTGGGCTGTGTCAGCATCCGTAATCTCTCTTCCG	840	
781	CCATTGTCGCGGCTGATCTTCTGGGCTGTGTCAGCATCCGTAATCTCTCTTCCG	840	
841	AAAGGCTCTAAATGACAACTGATTTCTCTGTAATCTCTCTTGTGTGCGAGTAATGCA	900	
841	AAAGGCTCTAAATGACAACTGATTTCTCTGTAATCTCTCTTGTGTGCGAGTAATGCA	900	
901	GTCATTAATTTTGGGCTCGGGCGGCTCGTTCTTAATCTTAAGCCCTTACGTAATCA	960	
901	GTCATTAATTTTGGGCTCGGGCGGCTCGTTCTTAATCTTAAGCCCTTACGTAATCA	960	

961	CAATTTGTGGCAAAATGGCGATGTGATGCCAGCAAGAAATCCTTGCCATTTTGACCGCA	1020	
961	CAATTTGTGGCAAAATGGCGATGTGATGCCAGCAAGAAATCCTTGCCATTTTGACCGCA	1020	
1021	TCAAGTTTGGCAGCAATGCGATAGATCTGTAAGACTTAACTTTGCTCATTTGCCGTT	1080	
1021	TCAAGTTTGGCAGCAATGCGATAGATCTGTAAGACTTAACTTTGCTCATTTGCCGTT	1080	
1081	CGGATTAAGTGTGGCGCATCTTCTTGGCGGTGAGACGACCTTTGTGACGTTGGCGCT	1140	
1081	CGGATTAAGTGTGGCGCATCTTCTTGGCGGTGAGACGACCTTTGTGACGTTGGCGCT	1140	
1141	GCGACATCGTTTGTGTTGAGCTGGTGAATCTTTCTAAACTGCAATAAACAATAAT	1200	
1141	GCGACATCGTTTGTGTTGAGCTGGTGAATCTTTCTAAACTGCAATAAACAATAAT	1200	
1201	CCGATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGAACCTTGAAGAGCTTTAG	1260	
1201	CCGATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGAACCTTGAAGAGCTTTAG	1260	
1261	CAGACAGCGCT 1271		
1261	CAGACAGCGCT 1271		
RESULT 2	1271 bp	DNA	Linear PAT 18-DEC-2003
AR391956	Sequence 6 from patent US 6613545.		
LOCUS	AR391956		
DEFINITION	Sequence 6 from patent US 6613545.		
ACCESSION	AR391956		
VERSION	AR391956.1 GI:40115727		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1271)		
AUTHORS	Kemernech, N., Salim, H., Eggeling, L. and Pfeifferle, W.		
TITLE	Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof		
JOURNAL	Patent: US 6613545-A 6 02-SEP-2003;		
FEATURES	Location/Qualifiers		
source	1..1271		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	100.0%;	Score 1271;	DB 6; Length 1271;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1271;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
1	GGCGATCAATGGAATCTAGCTTCAATATATGCAACAATAGCTAGTGGTGGCGCAAC	60	
1	GGCGATCAATGGAATCTAGCTTCAATATATGCAACAATAGCTAGTGGTGGCGCAAC	60	
61	TGGCAAAAACCTACCCGGCAATTTGTGTGATGATTTAGTGTGCAAAAAACGCAAGAT	120	
61	TGGCAAAAACCTACCCGGCAATTTGTGTGATGATTTAGTGTGCAAAAAACGCAAGAT	120	
121	TCATTCAAGCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA	180	
121	TCATTCAAGCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA	180	
181	TGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGA	240	
181	TGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGA	240	
241	CCGATGATGATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGATGGGCAAC	300	
241	CCGATGATGATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGATGGGCAAC	300	
301	CCCACTGTTTCCGGCTGATTTTCGGCGGCTCCACCGAATGCTGGTCACTCCCTCGT	360	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 5387 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6

Perfect score: 1271

Sequence: 1 gcgcgcacatgcaatctagct.....agcctaagcagacagcgcct 1271

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sbs:*
27: em_un:*
28: em_vl:*
29: em_hcg_hum:*
30: em_hcg_inv:*
31: em_hcg_other:*
32: em_hcg_mus:*
33: em_hcg_pin:*
34: em_hcg_rnd:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1271	100.0	1271	6 AR391953
2	1271	100.0	1271	6 AR391956
3	1271	100.0	1271	6 AX137709
4	1271	100.0	1271	6 AX137714
5	1271	100.0	1271	6 BD014991
6	1271	100.0	1271	6 BD014994
7	1271	100.0	340000	1 AP005274
8	1271	100.0	349980	6 AP120085
9	1260	99.1	2105	1 AP454053
10	753	59.2	753	6 AR391954
11	753	59.2	753	6 AX120373
12	753	59.2	753	6 AX137710
13	753	59.2	753	6 BD014992
14	753	59.2	753	6 BD162490
15	365.6	28.8	397	6 AX137083
16	365.6	28.8	778	6 AX137085
17	324	25.5	324	6 AR391955
18	324	25.5	324	6 AX120374
19	324	25.5	324	6 AX137712
20	324	25.5	324	6 BD014993
21	324	25.5	324	6 BD162491
22	185	14.6	715	6 AX137077
23	185	14.6	715	6 BD013995
24	149.6	11.8	87340	1 AP005224
25	132.8	10.4	53793	1 BX248361
26	127.4	10.0	711	6 AX123492
27	127.4	10.0	711	6 BD165609
28	127.4	10.0	309400	6 AX127153
29	127.4	10.0	325651	1 AP005283
30	59.8	4.7	349926	1 BX571660
31	58.6	4.6	756	6 AR396287
32	57.2	4.5	9888	1 AE011050
33	54.4	4.3	856	6 HSA335448
34	52.4	4.1	7218	6 166494
35	51.8	4.1	11025	1 AE013585
36	51.2	4.0	125020	9 AF429315
37	51	4.0	308015	1 AE016783
38	48.2	3.8	2000	6 AX655393
39	48	3.8	10732	6 E32986
40	46.6	3.7	3619	1 ECH410307
41	46.6	3.7	10733	1 AE013695
42	46.6	3.7	220050	1 AJ414156
43	46	3.6	215050	1 AL646084
44	45.8	3.6	20342	1 AE008888

ALIGNMENTS

RESULT 1	LOCUS	SEQUENCE	1 from patent US 6613545.	DNA	linear	PAT 18-DEC-2003
AR391953	AR391953	Sequence 1 from patent US 6613545.				
DEFINITION	AR391953	AR391953.1	GI:40115724			
ACCESSION	AR391953					
VERSION	AR391953.1					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1271)					
AUTHORS	Kennetknecht,N., Salm,H., Eggeling,U. and Pfeifferle,M.					
TITLE	Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof					
JOURNAL	Patent: US 6613545-A 1 02-SEP-2003;					

RESULT 15

08TKG9 PRELIMINARY; PRT; 113 AA.
 ID 08TKG9;
 AC 08TKG9;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Branched chain amino acid transport protein Azid.
 GN MA3438.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 NX NCBI_TaxID=2214;
 RN (1)
 RC SEQUENCE FROM N.A.
 RP STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=1932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smitrov S., Arnoor D., Brown A.,
 RA Allen N., Naylor J., Strange-Thomann N., DeZellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tittell A., Ye W.,
 RA Zimmer A., Barber R.D., Cam I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Larder E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AE011050; AAM06805.1;
 DR InterPro; IPR008407; Azid.
 DR Pfam; PF05437; Azid; 1.
 KW Complete proteome.
 SQ SEQUENCE 113 AA; 12626 MW; 90CBE128F935D7D CRC64;

Query Match 15.6%; Score 82; DB 17; Length 113;
 Best Local Similarity 28.6%; Pred. No. 1;
 Matches 32; Conservative 20; Mismatches 54; Indels 6; Gaps 4;

QY 1 MTTDESCILIVAVCAVITPALRAVPFLILKPRBSQFVGKAMMMFAGILAITASTFR 60
 DB 1 MMEPTLOMATIVIALATFATFTRVLPFCGRSREPPAMLSITEKNLPMILLLVYCYCK 60
 QY 61 SNAIDKTLTFGLIAY-AITVAHL-LGGRTLLSYGAGTIVFGIV--NLF 108
 DB 61 D--VQWFLAPYGIPELFTIGIVAGILMKRNMLSTFAGTGLYMLVQFNVF 110

Search completed: April 16, 2004, 06:55:52
 Job time : 23.3593 secs

DR Pfam: PF00211; guanylate_cyc; 2.
 DR SMART: SM00044; CYC; 2.
 DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 2.
 DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 2.
 DR Lyase.
 FT NON_TER 1 952 1
 FT NON_TER 1 952 1
 FT NON_TER 1 952 1
 SQ SEQUENCE 952 AA; 106967 MW; C7F30C7A67E6C5BB CRC64;
 Query Match
 Best Local Similarity 31.4%; Score 86; DB 11; Length 952;
 Matches 33; Conservative 19; Mismatches 31; Indels 22; Gaps 7;
 QY 18 ITFALRAVPELILKP---LRESQF---VGRMMMPAGILAILTASTFSSNAI---DLK 67
 DB 31 LTFALACFPALGAPALQEAQWARTVAEGWQ-----LTVTFVSVALPVSIL 83
 QY 68 TLTFGLIAVA--ITVVAHLGGRTLL--SVGAGTIVFGLVNL 108
 DB 84 AIGFGLVVAASHLVATLVPAKPRPRLWRTLGANLLFVG-VNMY 127
 RESULT 9
 Q8NFMS PRELIMINARY; PRT; 301 AA.
 ID Q8NFMS
 AC Q8NFMS
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Adenylate cyclase type I (EC 4.6.1.1) (Fragment).
 GN AOCY1
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22391367; PubMed=12503609;
 RA Ludwig M.G., Seewen K.;
 RT "Characterization of the human adenylate cyclase gene family: cDNA,
 RT gene structure, and tissue distribution of the nine isoforms."
 RL J. Recept. Signal Transduct. Res. 22:79-110(2002).
 DR EMBL: AF497515; AAM94372.1;
 DR GO: GO:0008294; F:calcium/calmodulin-responsive adenylate cyc. .; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 KM Lyase
 FT NON_TER 301 301
 FT NON_TER 301 AA; 32062 MW; 6AB95F1B98F466C6 CRC64;
 SQ SEQUENCE
 Query Match
 Best Local Similarity 31.4%; Score 85; DB 4; Length 301;
 Matches 33; Conservative 18; Mismatches 44; Indels 10; Gaps 5;
 QY 12 VAVCAVITFARAVPELILKRLRESQFVGKAMMPGILAILTASTFSSNAI---DLK 67
 DB 125 LALPFLFALCPFALGAPARSAAGAPARABEGWOLIV-TFVSVALPVSIL 103
 QY 68 TLTFGLIAVA--ITVVAHLGGRTLL--SVGAGTIVFGLVNL 108
 DB 184 AIGFGLVVAASHLVATLVPAKPRPRLWRTLGANLLFVG-VNMY 227
 RESULT 10
 Q92BR3 PRELIMINARY; PRT; 108 AA.
 ID Q92BR3
 AC Q92BR3
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein lin1481.
 GN LIN1481.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1642;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CIP 1162 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangoul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Bioeche H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
 RA Entian K.-D., Feih H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft U., Kuhn M., Kunst F., Kurapkhat G.,
 RA Madueno E., Matouman A., Mala Vicente J., Ng E., Nedjati H.,
 RA Nordstiek B., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Valquez-Boland J.-A., Vose H., Weiland J., Cossart P.;
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL596168; CAC96712.1; -
 DR PIR: AH1617; AH1617.
 DR Listlist; LIN01481;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 108 AA; 12014 MW; 89B14D6C6B8E083 CRC64;
 Query Match
 Best Local Similarity 24.8%; Score 84.5; DB 16; Length 108;
 Matches 25; Conservative 24; Mismatches 43; Indels 9; Gaps 3;
 QY 9 LTVAVCAVITFARAVPELILKRLRESQFVGKAMMPAGILAILTASTFSSNAIDLK 67
 DB 8 LTVLGGVLFVFRVLPFVRKQLQPSVIRLSVIRLSCILTFALFVOSLLITNSFP 67
 QY 68 TLTFGLIAVAITVVAHLGGRTLLSVGAGTIVFGLVNL 107
 DB 68 SINENMLATPILIT-----ALTKNLMIVIVGISM 101
 RESULT 11
 Q7VI97 PRELIMINARY; PRT; 113 AA.
 ID Q7VI97
 AC Q7VI97
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN HH0711.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Helicobacteraceae; Helicobacter.
 OC NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 RX MEDLINE=22709201; PubMed=12810954;
 RX Suerbaum S., Josephans C., Stenzenbach T., Dreescher B., Brandt P.,
 RA Bell M., Droege M., Partmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koening J., Macko L., Mendz G.L., Nyakatura G.,
 RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
 RT "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL: AE017146; AAP77308.1; -
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 113 AA; 12511 MW; 6D613F77CEB6C90F CRC64;
 Query Match
 Best Local Similarity 27.0%; Score 84.5; DB 16; Length 113;
 Matches 27; Conservative 27; Mismatches 39; Indels 7; Gaps 4;
 QY 9 LTVAVCAVITFARAVPELILKRLRESQFVGKAMMPAGILAILTASTFSSNAIDLK 68
 DB 9 VLIALIALNTLSRLPILIFAKSTPT-FIVSLGKVPISALIMILIVCLKD--MDLEH 65

RT avermitilins: deducing the ability of producing secondary
metabolites";
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22608306; Pubmed=12692562;
RX Ikeda H., Ishikawa J., Hamano A., Shimono M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Obara S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis".
RT Nat. Biotechnol. 21:526-531(2003).
RL EMBL, AP005031; BAC70208.1; -
RW Hypothetical protein; Complete proteome.
SQ SEQUENCE 102 AA; 10210 MW; 1D803E3DA40563B CRC64;

	Query Match	17.1%;	Score	90;	DB	16;	Length	102;	
	Best Local Similarity	29.7%;	Pred.	No. 0.14					
Matches	30;	Conservative	18;	Mismatches	43;	Indels	10;	Gaps	3;
QY	8	ILLVAVCVITVEFALRAVPFLIKLPRESQFGKAMMPAGIILATLTA-STF---RSNA	63						
		::::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::							
Dd	3	IWIAIAARVGGYAKVLGLVPAGVLBERPVRLLAALLPVALLATTAOCTFADGRITLV	62						
QY	64	IDKLTLTFLGLAVLTVAHHLGGRRTLSVGAGIIIVFGL	104						
		: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::							
Dd	63	LDAAKAAGLAANAVALTRAPF-----LLVAAAVVTNGV	97						

```

RESULT 6
08G3T6
ID      08G3T6          PRELIMINARY;          PRT;          110 AA.
AC      08G3T6;
DT      01-MAR-2003 (TREMBLrel.. 23, Created)
DT      01-MAR-2003 (TREMBLrel.. 23, Last sequence update)
DT      01-OCT-2003 (TREMBLrel.. 25, Last annotation update)
DE      Branched-chain amino acid permease.
DE      ALZD OR BL1668.
OC      Bifidobacterium longum.
OC      Bacteria: Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC      Bifidobacteriaceae; Bifidobacterium.
OX      NCBI_TaxID=216816;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NCC 2705;
RX      MEDLINE=22294977; PubMed=12381787;
RA      Schell M.A., Karamitranzou M., Stel B., Vilanova D., Berger B.,
RA      Pessi G., Zvaiflen M.-C., Destiere F., Bork P., Delley M.,
RA      Prichore R.D., Arigoni F.;
RT      "The genome sequence of Bifidobacterium longum reflects its adaptation
RT      to the human gastrointestinal tract.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR      EMBL; AE014801; AAN25455.1; -.
DR      InterPro; IPR008407; AZLD.
DR      Pfam; PF05437; AzLD; 1.
KW      Complete proteome.
SQ      SEQUENCE   110 AA;  12265 MW;  52FB38CF95274CC CRC64;

Query Match           16.8%; Score 88.5; DB 16; Length 110;
Best local Similarity 30.7%; Pred. No. 0.22;
Matches    35; Conservative 19; Mismatches 45; Indels 15; Gaps 6

QY      1 MTTFDFSCILVAVCAVITFALRAVPFLID----KPLRESQFYGXKAMMMNPAIGILAITA 56
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      3 MTWVGAGLIVTVVLG--TWCTRFLPFLVPESSQPRIIEYGV---LPVATGLLV 57

QY      57 STRSNAIDLKTLTFG--LIAVAITYAHLGGRRLLTSVGACTIYFVGNLT 107
       :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB      58 YALRNTPI--LTGSHGFLPIACTIVYLH-VMKRNMLLSIAGSLIVYMVLQL 108

```

```

ID      OBP824;      PRELIMINARY;      PRT;      112 AA.
AC
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE      Branched-chain amino acid transport protein.
GN      AZLD OR MM33263.
OS      Methanosarcina mazei (Methanosarcina frisia).
OC      Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC      Methanosarcinales; Methanosarcinaceae; Methanosarcina.
NCBI    TaxID=2209;
RN
RP      [1]
RP      SEQUENCE FROM N.A.
RC      STXAIN=Goel / GCI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX      MEDLINE=22120827; PubMed=12125824;
RA      Deppe-meier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA      Mattheis-Artis R., Henne A., Wieser A., Baumer S., Jacobi C.,
RA      Brieglebmann H., Lienard T., Christmann A., Boesicke M., Steckel S.,
RA      Bhattacharya A., Lytkin A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA      Fritz H.-J., Gottschalk G.;
RT      "The genome of Methanosarcina mazei: evidence for lateral gene
RL      transfer between Bacteria and Archaea.";
DR      J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR      EMBL; AE013585; AAM32959.1; -.
DR      InterPro; IPR008407; AZLD.
DR      Pfam; PF05437; AZLD. 1.
KW      Complete proteome.
SQ      SEQUENCE 112 AA; 12433 MW; 33D73C78F8DDA65C CRC64;

```

```

Query Match Similarity      16.5%; Score 87; DB 17; Length 112;
Best Local Similarity       29.4%; Pred.No. 0.32;
Matches    30; Conservative   19; Mismatches    49; Indels     4; Gaps     3

Cy          8 ILVVAVCAVITFAIRANVPFLIKLRRESQFYGKXAMMPAGIAIILTASTRSNAIDLK 67
Db         :|::|||::|
Cy        7 MLVTATVALNFAFRVRPPLFCFGSRPPAMISTEXKNIPMLIVLYCYLKD--YQWF 64
Db         |::|||::|
           68 TLTFGLIAV-AITVAHL-LGGRRTLSVGACTIVEGVNL 107
           :|::|||::|
           65 SAPYGPPELFETIGVAVAGHFWKRSMAMLSTFACTGIYMALVOL 106
           :|::|||::|

RESULT 8
ID 088444 PRELMINARY; PRT; 952 AA.
AC 088444;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Adenylyl cyclase type I (EC 4.6.1.1) (Fragment).
GN ADCYL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mamalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxId=10090; [1]
RX RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=98324784; PubMed=9662407;
RA Abdel-Majid R.M., Leong W.L., Schalkwyk L.C., Smallman D.S.,
RA Wong S.T., Storm D.R., Fine A., Dobson M.J., Guernsey D.L.,
RA Neumann P.E.;
PT "Loss of adenylyl cyclase I activity disrupts patterning of mouse
RL somatosensory cortex.";
RL Nat. Genet. 19:289-291(1998) .
DB EMBL; AF051880; AAC29478.1; -
DR HSSP; P26769; IAB8.
DR MOD; MG199677; ADcyl.
DR GO; GO:0008954; Focalin/caldonulin-responsive adenylate cyc.-.; IEA.
DR GO; GO:0004383; Figuanilate cyclase activity; IEA.
DR GO; GO:0016829; Filiyase activity; IEA.
DR GO; GO:0007642; P-intracellular signaling cascade; IEA.
DR InterPro; IPRO01054; G_cyclase.
```

DB 1 MTDFSCILLVAVCAVITPALRAVPLIKPRESQFVKMMMPAGIILASTFR 60
 61 SNAIDKTLTFFGLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 108
 DB 61 SNAIDKTLTFFGLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 108

RESULT 2

Q8NL67 PRELIMINARY; PRT; 115 AA.
 AC Q8NL67;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical membrane protein Cg13082.
 GN Cg13082.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005283; BAC00476.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID.1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 115 AA; 12018 MW; D18F27D197347F7D CRC64;

Query Match 24.7%; Score 130; DB 16; Length 115;
 Best Local Similarity 35.0%; Pred. No. 1.6e-05;
 Matches 36; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

QY 8 ILVVAV---CAVITPALRAVPLIKPRESQFVKMMMPAGIILASTFRSNAI 64
 DB 12 LNVAAVLTIPALITLLRIFPAAMKRVNSNOLMGVLRITPVGVVTVLITFGVYS 71
 QY 65 DLKLTITGLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 107
 DB 72 APGVGASLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 113

RESULT 3

Q8FSW9 PRELIMINARY; PRT; 116 AA.
 AC Q8FSW9;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN CE2925.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Ueda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005224; BAC19735.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID.1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 116 AA; 12255 MW; 31B0AAE454103417 CRC64;

Query Match 24.4%; Score 128.5; DB 16; Length 116;
 Best Local Similarity 30.7%; Pred. No. 2.2e-05;
 Matches 31; Conservative 27; Mismatches 40; Indels 3; Gaps 2;

QY 8 ILVVAVCAVITPALRAVPLIKPRESQFVKMMMPAGIILASTFRSNAIDK 67
 DB 15 VAVLIPVAILITLRLFPFVAMRGVKNQLMGVLSRTPGVVSVLVITLFTSTGPG 74
 QY 68 TLTFGLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 107
 DB 75 GIWASLIAVITLILH--WMRSAGLSIVGTVAIVAVLVNV 113

RESULT 4

Q9CF69 PRELIMINARY; PRT; 108 AA.
 AC Q9CF69;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Unknown protein.
 GN Y0RC OR L1612.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jajllon O., Malarme K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.,
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 DR EMBL; AE006391; AAK05710.1; -
 DR PIR; D86826; D86826.
 KW Complete proteome.
 SQ SEQUENCE 108 AA; 12024 MW; 9EE4838C1DFCF972 CRC64;

Query Match 18.0%; Score 94.5; DB 16; Length 108;
 Best Local Similarity 29.9%; Pred. No. 0.054;
 Matches 32; Conservative 24; Mismatches 42; Indels 9; Gaps 4;

QY 3 TDPSCLLVAVCAVITPALRAVPLIKPRESQFVKMMMPAGIILASTFRS 61
 DB 2 SSFESITLIGALITVWISRLVPLIKKQSLPQIVVEYISFVPMVMSLWISNLF 61
 QY 62 NAIDKTLTFFGLIAVAITVAHLLGGRRTLSVGAGTIVFGLVNL 107
 DB 62 HLGHLPSVNMNNDLASIPVLAAL--TKNLL---VIVLVGFSL 101

RESULT 5

Q82XA6 PRELIMINARY; PRT; 102 AA.
 AC Q82XA6;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN SAV2497.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces

Tue Apr 20 06:47:22 2004

us-10-608-504-5.rpt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 21.3593 Seconds
(without alignments)
1595.367 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526

Sequence: 1 MTTDFSCILVAVACAVITF.....RTLSVAGACTIVFGLVLP 108

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526	100.0	108	16	Q8NTP7 corynebacte
2	130	24.7	115	16	Q8NTL67 corynebacte
3	128.5	24.4	116	16	Q8FSW9 corynebacte
4	94.5	18.0	108	16	Q9CF69 lactococcus
5	90	17.1	102	16	Q82KA6 streptococcus
6	88.5	16.8	110	16	Q8G3T6 bifidobacte
7	87	16.5	112	17	Q8PS24 methanosaarc
8	86	16.3	952	11	Q88444 mus musculu
9	85	16.2	301	4	Q8NFM5 homin sapien
10	84.5	16.1	108	16	Q92BR3 lactobacte
11	84.5	16.1	113	16	Q7V197 heliobacte
12	83.5	15.9	407	16	Q9A5Z6 caulobacte
13	83	15.8	108	16	Q8Y760 lacteria mo
14	82.5	15.7	388	16	Q8FIC1 escherichia
15	82	15.6	113	17	Q8TKG9 mechanosarc
16	81.5	15.5	109	16	Q8CU94 stephilotococ

17	81.5	15.5	182	16	Q81YS6 bacillus an
18	80.5	15.3	108	16	Q8BY20 lactobacilli
19	80.5	15.3	192	16	Q985A1 rhizobium l
20	80.5	15.3	345	16	Q9ZUW5 rhizobium
21	80	15.2	589	16	Q9PEE7 xyloella fas
22	80	15.2	589	16	Q87EFO xyloella fas
23	79.5	15.1	956	16	Q7U948 synecococc
24	79	15.0	148	16	Q8XDT4 escherichia
25	79	15.0	314	17	Q9YXE7 pyrococcus
26	79	15.0	394	16	Q8ZJZ6 salmoneilla
27	79	15.0	394	16	Q8ZJZ6 salmoneilla
28	78.5	14.9	388	16	Q8XA63 escherichia
29	78	14.8	443	16	Q98FMO rhizobium l
30	77.5	14.7	235	16	Q45Z35 bradyrhizob
31	77	14.6	388	16	Q83SR3 shigella fl
32	77	14.6	464	16	Q8DWD1 streptococc
33	76.5	14.5	348	16	Q8XDV2 chlorobium
34	76.5	14.5	440	16	Q8NMG2 corynebacte
35	76	14.4	102	16	Q50482 streptococ
36	76	14.4	114	16	Q7MMJ7 bordetella
37	76	14.4	114	16	Q7MB31 bordetella
38	76	14.4	402	16	Q89RX2 bradyrhizob
39	76	14.4	459	16	Q88BE1 pseudomonas
40	76	14.4	543	3	Q86ZM8 podospira a
41	76	14.4	656	16	Q9KRX4 vibrio chol
42	75.5	14.4	107	16	Q8DVS8 streptococ
43	75.5	14.4	182	17	Q9YC04 aeropyrum p
44	75.5	14.4	506	13	Q8UVB7 brachydanio
45	75.5	14.4	506	13	Q7ZU13 brachydanio

ALIGNMENTS

ID	Q8NTP7	PRELIMINARY;	PRT;	108 AA.
AC	Q8NTP7	01-OCT-2002 (TREMURel. 22, Created)		
DT	01-OCT-2002 (TREMURel. 22, Last sequence update)			
DT	01-OCT-2003 (TREMURel. 25, Last annotation update)			
DE	Hypothetical membrane protein Cg10259 (BrnE).			
GN	Cg10259 OR BRNE.			
OC	Corynebacterium glutamicum (Brevibacterium flavum).			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.			
OX	NCBI_TaxID=1718;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;			
RA	Nakagawa S.,			
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."			
RL	Submitted (May-2002) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22077265; PubMed=12081967;			
RA	Kennedynecht N., Sahm H., Yen M.R., Patek M., Sailer M.H. Jr.,			
RA	Eggeling L.,			
RT	"Export of L-Isoleucine from Corynebacterium glutamicum: a Two-Gene-			
RT	Encoded Member of a New Translocator Family."			
RL	J. Bacteriol. 184:3947-3956(2002).			
DR	EMBL; AP005274; BAB97652.1; -			
DR	EMBL; AF454053; AAM4685.1; -			
DR	InterPro: IPR008407; AcLD.			
DR	Pfam: PF05437; AcLD.			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 108 Aa; 11480 MW; 3A160B08422CC21PD CRC64;			

Query Match 100.0%; Score 526; DB 16; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.7e-45;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MTTDFSCILVAVACAVITFALRAVPLILKPLRESQFVGGMAMMAGILALITASTFR 60

FT DOMAIN 92 96 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 97 117 3 (POTENTIAL).
 FT DOMAIN 118 127 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 128 148 4 (POTENTIAL).
 FT DOMAIN 149 156 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 157 177 5 (POTENTIAL).
 FT DOMAIN 178 182 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 183 203 6 (POTENTIAL).
 FT DOMAIN 204 257 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 258 278 7 (POTENTIAL).
 FT DOMAIN 279 293 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 294 314 8 (POTENTIAL).
 FT DOMAIN 315 320 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 321 341 9 (POTENTIAL).
 FT DOMAIN 342 367 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 368 388 10 (POTENTIAL).
 FT DOMAIN 389 404 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 405 425 11 (POTENTIAL).
 FT DOMAIN 426 438 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 439 459 12 (POTENTIAL).
 FT DOMAIN 460 477 DILEUCINE INTERNALIZATION MOTIF (BY
 FT SITE 12 13 SIMILARITY).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (By similarity).
 FT CONFLICT 39 39 S -> N (IN REF. 1).
 FT CONFLICT 94 94 S -> A (IN REF. 2 AND 4).
 FT CONFLICT 429 429 S -> N (IN REF. 1).
 FT CONFLICT 477 AA; 51523 MW; A3753FB34545F9A CRC64;
 SQ SEQUENCE

Query Match 13.3%; Score 70; DB 1; Length 477;
 Best Local Similarity 31.8%; Pred. No. 10;
 Matches 27; Conservative 13; Mismatches 27; Indels 18; Gaps 4;

QY 29 ILKPLRESQVFGKMMNAGILALITASTERSNAI-----DLKTLFGILAVIT 79
 DB 254 IKKPLIG--ISLMPQQLSGVNAIM---FYANSIFEBAKKDSLSAVIGIQLVLT 307
 QY 80 VVAHLT---GGRRTLISVGATIVF 101
 DB 308 AVAALIMDRAGRRLILALSGVIMVF 332

RESULT 15
 NUSM PODAN STANDARD; PRT; 652 AA.
 ID NUSM PODAN
 AC P20679;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
 GN ND5.
 OS Podospora anserina.
 OG Mitochondrion.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
 OC NCBI_TaxID=5145;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S, and A;
 RX MEDLINE=90204555; Pubmed=2319602;
 RA Cummings D.J., Michel F., Domenico J.M., McNally K.L.;
 RT "DNA sequence analysis of the mitochondrial ND4L-ND5 gene complex
 RT from Podospora anserina. Duplication of the ND4L gene within its
 RT intron".
 RT J. Mol. Biol. 212:269-286(1990).
 RL [2]
 RP COMPLETE GENOME.
 RC STRAIN=S;
 RX MEDLINE=90291512; Pubmed=2357736;
 RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;
 RT "The complete DNA sequence of the mitochondrial genome of Podospora
 RT anserina".
 RL Curr. Genet. 17:375-402(1990).

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC
 CC EMBL; X55026; CA38798.1; -.
 DR PIR; S09133; S09133.
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR003916; NADH_oxred5.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1; 1.
 DR Pfam; PF00662; Oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADH_OXRED5.
 DR PRINTS; PR01437; NUOXDRDASE4.
 DR OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.
 KM SEQUENCE 652 AA; 72632 MW; 84CA87B75755F9A CRC64;
 SQ

Query Match 13.3%; Score 70; DB 1; Length 652;
 Best Local Similarity 25.9%; Pred. No. 13;
 Matches 28; Conservative 19; Mismatches 37; Indels 24; Gaps 3;

QY 6 SCILVAVCAVTFALRAVPFILKPLRESQVFG-----KMMNMPAGILALITASTER 60
 DB 206 SKVLLIGIC-----LLIGAMAKSSQVGLHWLPMAMBGPPTVPSALIHATWV 253
 QY 61 SNADLKTTLFGILAVITVAHLGGRRTLISVGATIVFGLVNLV 108
 DB 254 TAGVYLIMRSSPLIEYNSVTL-----LCWLGAITTVFSSLIQFL 294

Search completed: April 16, 2004, 06:53:45
 Job time : 7.01671 secs

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF084229; AAD52041.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.

DR PRINTS: PF00001; 7tm.1; 1

DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 32
 FT TRANSMEM 33 59
 FT TRANSMEM 60 68
 FT TRANSMEM 69 88
 FT TRANSMEM 90 103
 FT TRANSMEM 104 125
 FT TRANSMEM 126 143
 FT TRANSMEM 144 164
 FT TRANSMEM 165 187
 FT TRANSMEM 188 215
 FT TRANSMEM 216 231
 FT TRANSMEM 232 259
 FT TRANSMEM 260 275
 FT TRANSMEM 276 293
 FT TRANSMEM 294 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

FT TRANSMEM 342 AA; 39273 MW; 585800397806622A CRC64; N-LINKED (GLCNAC...) (POTENTIAL).

RT glucose transport activity." J. Biol. Chem. 275:16275-16280 (2000).
 RL J. Biol. Chem. 275:16275-16280 (2000).
 RN (3)
 RC SEQUENCE FROM N.A.
 RA STRAIN=129; Tissue=Embryonic carcinoma;
 RX MEDLINE=201319023; PubMed=10660996;
 RA Carayannopoulos M.O.; Chi M.M.-Y.; Cui Y.; Pingrethaus J.M.;
 RA McKnight R.A.; Mueckler M.; Devaskar S.U.; Moley K.H.;
 RT "Glucose is a glucose transporter responsible for insulin-stimulated
 RT glucose uptake in the blastocyst."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318 (2000).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Ola; Tissue=Spleen;
 RX MEDLINE=21547794; PubMed=11689004;
 RA Scheepers A.; Doege H.; Joost H.-G.; Schnermann A.;
 RT "Mouse GLUT8, genomic organization and regulation of expression in
 RT 3T3-L1 adipocytes by glucose."
 RL Biochem. Biophys. Res. Commun. 288:969-974 (2001).
 CC -1- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration (By similarity). Insulin induces a
 CC change in the intracellular localization and gives rise to
 CC insertion in the plasma membrane.
 CC -1- TISSUE SPECIFICITY: Highest level of expression in placenta and
 CC testis. Highly expressed in adult and pubertal testis, but not
 CC prepubertal testis. Lower levels of expression in brain, liver,
 CC heart, kidney, fat and skeletal muscle.
 CC -1- DEVELOPMENTAL STAGE: High expression in blastocysts.
 CC -1- INDUCTION: Inhibited under glucose deprivation.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTERS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ245936; CAB75719.1; -
 DR EMBL: Y17802; CAB89815.1; -
 DR EMBL: AF232061; AAF78366.1; -
 DR EMBL: AJ413951; CAC98690.1; -

DR WGD: WGI:1860103; SLC2A8.
 DR GO: GO:0005887; C:intracellular to plasma membrane; IDA.
 DR GO: GO:0005536; F:glucose binding; IDA.
 DR GO: GO:0005355; F:glucose transporter activity; IDA.
 DR GO: GO:0015758; P:glucose transport; IDA.
 DR GO: GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO: GO:0001665; P:response to hypoxia; IDA.

DR InterPro: IPR007114; WFS.
 DR InterPro: IPR005828; SUGAR_TRANSPORT_1; 1.
 DR InterPro: IPR005829; SUGAR_TRANSPORT_2; 1.
 DR InterPro: IPR003663; SUGAR_TRANSP.

DR Pfam: PF00083; sugar tr; 1.
 DR PRINTS: PR00171; SUGRTRANSPORT.
 DR TIGRfam: TIGR006879; SP: 1.

DR PROSITE: PS00850; MFS; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.

DR TRANSMEM 1 25
 DR TRANSMEM 26 46
 DR TRANSMEM 47 70
 DR TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

FT TRANSMEM 1 25
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

RT immunodeficiency viruses.";
 RL Nature 388:296-300(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT ALA-25.
 RX MEDLINE=97311099; PubMed=9166430;
 RA Liao P., Alkhatib G., Peden K.W.C., Sharma G., Berger E.A.,
 RA Fader J.M.,
 RA "STRL3, a novel chemokine receptor-like protein, functions as a
 RT fusion cofactor for both macrophage-tropic and T cell line-tropic
 RT HIV-1.";
 RL J. Exp. Med. 185:2015-2023(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Blood;
 RX MEDLINE=97431687; PubMed=9285716;
 RA Loetscher M., Amara A., Oberlin E., Brass N., Legler D.F.,
 RA Loetscher P., D'Apuzzo M., Meese E.U., Rouselet D., Virelizier J.L.,
 RA Baggiolini M., Arenzana-Seisdedos F., Moser B.,
 RA "TYWSTR, a putative chemokine receptor selectively expressed in
 RT activated T cells, exhibits HIV-1 coreceptor function.";
 RL Curr. Biol. 7:652-660(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Kopatz S.A., Aronstam R.S., Sharma S.V.,
 RT "cDNA clones of human proteins involved in signal transduction
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Alteschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.W., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Receptor for the C-X-C chemokine CXCL6. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in lymphoid tissues and activated T
 CC cells.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF007545; AAB64221.1; -
 CC DR EMBL; U73529; AAB61456.1; -
 CC DR EMBL; U73531; AAB61457.1; -
 CC DR EMBL; Y13248; CA473698.1; -
 CC DR EMBL; AY322543; AAB84356.1; -
 CC DR EMBL; BC033584; AAB33584.1; -
 CC DR Genew; HGNC:16647; CXCR6.

DR MIM; 605163; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0015026; F:coreceptor activity; TAS.
 DR GO; GO:0004930; F:G-protein coupled receptor activity; TAS.
 DR GO; GO:0007186; F:G-protein coupled receptor protein signalin...; TAS.
 DR GO; GO:0008166; P:viral replication; TAS.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOSPN.
 DR PROSITE; PS00252; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
 FT DOMAIN 1
 FT TRANSMEM 33 59
 FT DOMAIN 2
 FT TRANSMEM 60 68
 FT TRANSMEM 69 89
 FT TRANSMEM 90 103
 FT TRANSMEM 104 125
 FT DOMAIN 3
 FT TRANSMEM 126 143
 FT TRANSMEM 144 164
 FT TRANSMEM 165 187
 FT TRANSMEM 188 215
 FT TRANSMEM 216 231
 FT TRANSMEM 232 259
 FT TRANSMEM 260 275
 FT TRANSMEM 276 293
 FT DOMAIN 4
 FT TRANSMEM 294 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16
 FT VARIANT 25 25
 FT SO
 SEQUENCE 342 AA; 39280 MM; 9FBC0255601082E CRC64;
 Query Match 13.3%; Score 70; DB 1; Length 342;
 Best Local Similarity 27.8%; Pred. No. 7.3;
 Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;
 QY 7 CILIVAVCAVITFAFAV-----FLIKPAREQFVQKAMMPAGILAI 53
 DB 36 CMYLVAVVCGVGNGLVIVISYHQLQSTIDVFLVPLADLVFCTLPFMYAGIHEW 95
 QY 54 LTATFRSNADIKTLTF---GLIYAVITV 80
 DB 96 VFGQWCKSLIGIYITINFTYMLITCIV 125
 RESULT 13
 ID CCR6_PANTR STANDARD; PRT; 342 AA.
 AC Q9TVI6;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C Chemokine receptor, type 6 (CXCR-6) (G protein-coupled
 DE receptor bonzo) (G protein-coupled receptor STRL33).
 GN CXCR6 OR BONZO OR STRL33.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCB:Taxid=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99433499; PubMed=10505680;
 RA Brüssel A., Pretet J.-L., Girard M., Butor C.,
 RT "Sequences and predicted structures of chimpanzee STRL33 (Bonzo) and
 RT gR15 (BOB).";
 RT AIDS Res. Hum. Retroviruses 15:1315-1319(1999).
 CC -1- FUNCTION: Receptor for the C-X-C chemokine CXCL6. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

DR PIR; S04254; G31277.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub_transporter.
 DR InterPro; IPR005829; Sug_transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR TIGRfam; TIGR00879; Sp; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane; Transport; Quinate metabolism; Glycoprotein.
 FT DOMAIN 1 26
 FT TRANSSEM 27 47
 FT TRANSSEM 48 74
 FT TRANSSEM 75 95
 FT TRANSSEM 96 98
 FT TRANSSEM 99 119
 FT TRANSSEM 120 131
 FT TRANSSEM 132 152
 FT TRANSSEM 153 160
 FT TRANSSEM 161 181
 FT TRANSSEM 182 195
 FT TRANSSEM 196 216
 FT TRANSSEM 217 285
 FT TRANSSEM 286 306
 FT TRANSSEM 307 327
 FT TRANSSEM 328 349
 FT TRANSSEM 350 352
 FT TRANSSEM 353 373
 FT TRANSSEM 374 389
 FT TRANSSEM 390 410
 FT TRANSSEM 411 435
 FT TRANSSEM 436 456
 FT TRANSSEM 457 458
 FT TRANSSEM 459 479
 FT TRANSSEM 480 537
 FT CARBOHYD 184 184
 FT SEQUENCE 537 AA; 60103 MW; 9AC63400FC164F3 CRC64;
 Query Match 13.7%; Score 72; DB 1; Length 537;
 Best Local Similarity 35.4%; Pred. No. 7.1;
 Matches 35; Conservative 16; Mismatches 26; Indels 22; Gaps 9;
 QY 26 PFLIKPARESQ---FVGKKA-MWM-PAGILAI--LTASFRSNAID-----LKTLPFG 72
 DB 273 PFLSIKQ-RKQWRFIFGMLTFPGWNGSGINAINYSPTVFRSIGITGDTGFLTIGIFG 331
 QY 73 LIAVAIVTVVAHL---LGRRTLLISVGA--GTV--FVG 103
 DB 332 VVKVLTITLIMLMLVDLVRRIILFTGAAGSLDMWIFG 370
 RESULT 11
 CCR6_CERAE STANDARD; PRT; 342 AA.
 ID 016983;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled
 DE receptor bonzo).
 GN CCR6 OR BONZO.
 OS Cercopithecus aethiops (Green monkey) (Primate).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97373958; PubMed=9230441;
 RA Deng H.K., Unutmaz D., Kewalramani V.N., Littman D.R.,
 RT "Expression cloning of new receptors used by simian and human

RT Immunodeficiency viruses."
 RL Nature 388:296-300(1997).
 CC -!- FUNCTION: Receptor for the C-X-C chemokine CXCL16. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR ENBL; AF007859; AAB64225.1; ..
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHDOPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS02622; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 32
 FT TRANSSEM 33 39
 FT TRANSSEM 60 68
 FT TRANSSEM 69 89
 FT TRANSSEM 90 103
 FT TRANSSEM 104 125
 FT TRANSSEM 126 143
 FT TRANSSEM 144 164
 FT TRANSSEM 165 187
 FT TRANSSEM 188 215
 FT TRANSSEM 216 231
 FT TRANSSEM 232 259
 FT TRANSSEM 260 275
 FT TRANSSEM 276 293
 FT TRANSSEM 294 342
 FT TRANSSEM 342 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16
 FT SEQUENCE 342 AA; 39226 MW; 6CBFB389CE5919E CRC64;
 Query Match 13.3%; Score 70; DB 1; Length 342;
 Best Local Similarity 27.8%; Pred. No. 7.3;
 Matches 25; Conservative 11; Mismatches 38; Indels 16; Gaps 2;
 QY 7 CILVVAVCAVTFEFAFRAV-----PFLIKPARESQVGGKAMMPAGILAI 53
 DB 36 CMLIVFVCGVANSVLVVISFPHKLGSLTFVFNPLDVLVVCILPFMAVAGIHEN 95
 QY 54 LTASTFRSNAIDKLTTF--GLIAVAIVTV 80
 DB 96 IFGQVCKTLLGIYITINFTSMILITCIV 125
 RESULT 12
 CCR6_HUMAN STANDARD; PRT; 342 AA.
 ID 000574; 000575;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled
 DE receptor bonzo) (G protein-coupled receptor STR333).
 GN CCR6 OR BONZO OR STR333 OR TYM6TR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97373958; PubMed=9230441;
 RA Deng H.K., Unutmaz D., Kewalramani V.N., Littman D.R.,
 RT "Expression cloning of new receptors used by simian and human

```

FT DOMAIN 497 538 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 539 557 7 (BY SIMILARITY).
FT DOMAIN 558 572 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 873 892 8 (BY SIMILARITY).
FT DOMAIN 893 898 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 899 918 9 (BY SIMILARITY).
FT DOMAIN 919 924 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 925 946 10 (BY SIMILARITY).
FT DOMAIN 947 974 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 975 993 11 (BY SIMILARITY).
FT DOMAIN 994 1006 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 1007 1029 12 (BY SIMILARITY).
FT DOMAIN 1030 1049 CYTOPLASMIC (BY SIMILARITY).
SQ SEQUENCE 1049 AA; 113573 MW; 19670E3C4CC29055 CRC64;

Query Match 14.0%; Score 73.5; DB 1; Length 1049;
Best Local Similarity 19.4%; Pred. No. 9.5;
Matches 26; Conservative 22; Mismatches 33; Indels 53; Gaps 3;

QY 5 FSCILLVAVCAVITFALRAVPEFLIKFRESQVGRKAMMPAGIALITASTFR--- 60
DB 877 YAIISLVFELCLAAAYESMSIPFSVM-----LVVPLGVIGALLAATFRGLTN 923
QY 61 -----SNAI-----DKLTTFGLIAVATVVAHL 84
DB 924 DYFQVGVGLITTTGLSAKNILIVERPAKDMKEKGLIEATLDVAVRRLRLMTSLAFL 983
QY 85 IGGRTTLASVAGT 98
DB 984 LGVPLVLVISTAGS 997

RESULT 9
Y420_NEIMA STANDARD; PRT; 227 AA.
ID Y420_NEIMA
AC Q9JRI0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
GN Hypothetical protein NMA0420/NMB2020.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RC MEDLINE=20222556; PubMed=10761919;
RA Parhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holtroyd S.,
RA Jagels K., Leathers S., Moule S., Mungall K., Quail M.A.,
RA Rajadaram M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrett B.G.;
RA "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RT Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RC MEDLINE=2015755; PubMed=10710307;
RA Tettein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Kechum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gilm M.L., Debey R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., Parksey D.S., Blair B., Citterone H., Clark E.B.,
RA Mason T.D., Uetdeck T.R., Khouri H., Qin H., Vamthyan J.,
RA Gili J., Scarlato V., Maatigani V., Pizarro M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappaport J., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."

```

```

RL Science 287:1809-1815(2000).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AL162753; CAB83719.1; --
DR EMBL; AE002552; AAF42343.1; --
DR PIR; G81014; G81014.
DR TIGR; NMB2020; --
DR InterPro; IPR006214; UPF0005.
DR Pfam; PF01027; UPF0005; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
SQ SEQUENCE 227 AA; 24454 MW; 0E24A04F9676CDBD CRC64;

Query Match 13.7%; Score 72; DB 1; Length 227;
Best Local Similarity 31.9%; Pred. No. 3.2;
Matches 22; Conservative 14; Mismatches 21; Indels 12; Gaps 3;

QY 49 GILAIITASTF-----RSNAIDL---KTLTFGLIAVATVVAHL-IGGRTTLASVGA 96
DB 114 GIAAATRAVETMSALARRTRIDNNALGRFLTVGAVILMAVAVANLFGIPALALTISA 173
QY 97 GTIVFGLV 105
DB 174 GFVFLSSLM 182

RESULT 10
QAY_NEUCR STANDARD; PRT; 537 AA.
ID QAY_NEUCR
AC F11636;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Orotate permease (Orotate transporter).
GN QAY-.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RC MEDLINE=89293848; PubMed=2525625;
RA Geever R.F., Hulet L., Baum J.A., Tyler B.M., Patel V.B.,
RA Rutledge B.J., Case M.B., Giles N.H.;
RA "DNA sequence, organization and regulation of the qa gene cluster of
RT Neurospora crassa."
RT J. Mol. Biol. 207:15-34(1989).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the sugar transporter family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X14603; CAA32752.1; --

```

```

OX NCB1_TaxID=1488;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hittl J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- FUNCTION: Probably involved in tryptophan uptake (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE TRP TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL/AE007858; AAK81540.1; -.
CC
CC DR PIR; A97344; A97344.
CC KW Transport; Amino-acid transport; Transmembrane; Complete proteome.
CC FT TRANSMEM 9 31 POTENTIAL.
CC FT TRANSMEM 58 80 POTENTIAL.
CC FT TRANSMEM 107 129 POTENTIAL.
CC FT TRANSMEM 135 161 POTENTIAL.
CC SQ SEQUENCE 171 AA; 18440 MW; CC9AAB476CECB80C CR664;

Query Match 14.1%; Score 74; DB 1; Length 171;
Best local Similarity 22.8%; Pred. No. 1.5;
Matches 31; Conservative 23; Mismatches 44; Indels 38; Gaps 4;

QY 1 MTTDESCILVAV-----CAVTFALRAVP-----FL 28
DB 32 MKPDFSIAMLFITLLNDYKTCISTGVAGLAAAVTFPGGQDLNITDRVTTSLVFI 91
QY 29 ILKRESQFVGKAMMPAGILAITASTFRSNAIDKTLTFGLAVAVITVAHLGGR 88
DB 92 ALRPPFD-KINDKIHITITVGLTISGVFLGSALITVGLPSFALRITVVL-----P 145
QY 89 RFLSVAGATTIVGL 104
DB 146 ALTINAVGTILFVAV 161

RESULT 8
ACRB_ECOLI STANDARD; PRT; 1049 AA.
ID ACRB_ECOLI STANDARD; PRT; 1049 AA.
AC P31224;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acridiflavine resistance protein B.
GN ACRB OR ACRB OR B0462.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCB1_TaxID=562;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Xu U., Bertrand K.P.;
RT "Nucleotide sequence of the acraB operon from Escherichia coli.";
RL Submitted (May-1993) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / M4573;
RX MEDLINE=94012493; PubMed=8407802;

```

```

RA Ma D., Cook D.N., Alberti M., Pon N.G., Nikaide H., Hearst J.E.;
RT "Molecular cloning and characterization of acra and acrb genes of
RT Escherichia coli."
RL J. Bacteriol. 175:6299-6313(1993).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / M61655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1233-1238(1997).
RN
RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kallman S., Komp C., Kuri O.,
RA Lew H., Lin D., Narmath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN
RP CHARACTERIZATION.
RX MEDLINE=95379493; PubMed=7651136;
RA Ma D., Cook D.N., Alberti M., Pon N.G., Nikaide H., Hearst J.E.;
RT "Genes acra and acrb encode a stress-induced efflux system of
RT Escherichia coli."
RL Mol. Microbiol. 16:45-55(1995).
RN
RP [6]
RP INTERACTION WITH ACRA.
RX MEDLINE=20381028; PubMed=10920254;
RA Kawabe T., Fujihira E., Yamaguchi A.;
RT "Molecular construction of a multidrug exporter system, AcraB:
RT molecular interaction between Acra and Acrb, and cleavage of the
RT N-terminal signal sequence of Acra."
RL J. Biochem. 128:195-200(2000).
CC -1- FUNCTION: AcraB is a drug efflux protein with a broad substrate
CC specificity.
CC -1- SUBUNIT: INTERACTS WITH ACRA.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE ACRB/ACRD/ACRF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL/M94248; AAA23411.1; -.
CC DR EMBL; U00734; AAA67135.1; -.
CC DR EMBL; AE000152; AAC73564.1; -.
CC DR EMBL; U82664; AAB40216.1; -.
CC DR PIR; B36938; B36938.
CC Ecogene; BG11704; acrb.
CC DR InterPro; IPR001036; Acriflavin_res.
CC DR InterPro; IPR004764; HXBL.
CC DR Pfam; PF00873; ACR_tan1.
CC DR PRINTS; PR00702; ACRIFLAVINRP.
CC DR TIGRFAMs; TIGR00915; 2A0602; 1.
CC KW Transmembrane; Inner membrane; Transport; Complete proteome.
CC FT TRANSMEM 1 9
CC FT DOMAIN 10 28
CC FT DOMAIN 29 339
CC FT TRANSMEM 340 359
CC FT TRANSMEM 360 365
CC FT TRANSMEM 366 385
CC FT TRANSMEM 386 391
CC FT TRANSMEM 392 413
CC FT TRANSMEM 414 441
CC FT TRANSMEM 442 460
CC FT TRANSMEM 461 473
CC FT TRANSMEM 474 496
CC FT TRANSMEM 497 500

```

OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=9298382; PubMed=2472670;
RA Krupinski J., Coussen F., Bakalyar H.A., Tang W.-J.,
RA Feinstein P.G., Orth K., Slaughter C., Reed R.R., Gilman A.G.,
RT "adenyl cyclase amino acid sequence; possible channel- or
RT transporter-like structure."
RL Science 244:1558-1564(1989).
RN [2]
RP 3D-STRUCTURE MODELING OF 295-450; 861-936 AND 950-1045.
RX MEDLINE=98054247; PubMed=9391039;
RA Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.,
RT "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
RT and mutational analysis."
RL Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
CC -1- FUNCTION: This is a calmodulin-sensitive adenylyl cyclase. May be
CC involved in regulatory processes in the central nervous system. It
CC may play a role in memory acquisition and learning.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -1- ENZYME REGULATION: Activated by calcium/calmodulin. Inhibited by
CC the G protein beta and gamma subunit complex.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Brain.
CC -1- DOMAIN: Composed of two homologous domains.
CC -1- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl
CC cyclase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M25579; AAA79957.1; -
CC PIR: A41350; A41350.
CC PDB: 1AWK; 28-JAN-98.
CC InterPro: IPR001054; G_cyclase.
CC Pfam: PFO0211; guanylate_cyc; 2.
CC SMART: SMO0044; CYCG; 2.
CC PROSITE: PS00452; GUANYLATE CYCLASES_1; 2.
CC PROSITE: PS50125; GUANYLATE CYCLASES_2; 2.
CC Lyase; CAMP biosynthesis; Transmembrane; Glycoprotein; Repeat;
CC Calmodulin-binding; Metal-binding; Magnesium; 3D-structure.
KW DOMAIN 1
FT 1 65
FT TRANSMEM 66 86
FT TRANSMEM 90 110
FT TRANSMEM 127 147
FT TRANSMEM 154 174
FT TRANSMEM 184 204
FT TRANSMEM 216 236
FT DOMAIN 237 612
FT TRANSMEM 613 633
FT TRANSMEM 637 657
FT TRANSMEM 676 696
FT TRANSMEM 697 726
FT TRANSMEM 727 747
FT TRANSMEM 755 775
FT TRANSMEM 797 797
FT DOMAIN 798 1134
FT 495 522
FT DOMAIN 1027 1050
FT 3 34
FT METAL 310 310
FT METAL 311 311

FT METAL 354 354
FT CARBOHYD 706 706
FT STRAND 298 300
FT STRAND 304 310
FT STRAND 312 315
FT TURN 316 317
FT TURN 320 321
FT TURN 324 337
FT TURN 338 339
FT TURN 340 342
FT TURN 344 345
FT STRAND 348 348
FT STRAND 355 360
FT TURN 367 367
FT TURN 368 391
FT TURN 392 392
FT STRAND 398 404
FT STRAND 407 410
FT STRAND 418 419
FT STRAND 423 423
FT TURN 424 434
FT TURN 434 434
FT TURN 444 447
FT TURN 448 449
FT STRAND 863 864
FT STRAND 870 873
FT TURN 879 883
FT TURN 888 890
FT TURN 891 893
FT TURN 894 912
FT TURN 914 915
FT STRAND 924 925
FT TURN 926 927
FT STRAND 928 929
FT STRAND 931 933
FT TURN 951 952
FT TURN 953 971
FT TURN 972 973
FT STRAND 983 985
FT STRAND 990 992
FT STRAND 998 1000
FT TURN 1004 1014
FT TURN 1015 1015
FT TURN 1025 1030
FT TURN 1031 1033
FT TURN 1034 1035
SQ SEQUENCES 1134 AA; 123978 MW; CCA410BC224DF3 CRC64;
Query Match 14.3%; Score 75; DB 1; Length 1134;
Best Local Similarity 28.6%; Pred. No. 7.2;
Matches 30; Conservative 20; Mismatches 45; Indels 10; Gaps 5;
OY 12 VAVCAVTFALRAVPFILLPLEBSQVGVGMAMFAGILAILTASFPSNAI-----DLK 67
DB 127 LALLFSFTFALLCCFPALGSPAGAHAGAAVPAITADGVQMLLV-TFVSALLPVRSL 165
OY 68 TLTFGLIAVA--ITVVAHLIGSRITLL--SVGAGTIVFVGLVNL 108
DB 186 AIGGIVVAASHLVTATLVPAKRPRLRTTGAVALLFLG-VVVY 229
RESULT 7
TRPP_CLOAB STANDARD; PRT; 171 AA.
AC 097D63;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
GN Probable tryptophan transport protein.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.

FT SITE 225 225 IMPORTANT FOR PH SENSOR.
 SQ SEQUENCE 368 AA; 41355 MM; B508BD25EE9130 CRC64;
 Query Match 15.5%; Score 81.5; DB 1; Length 388;
 Best Local Similarity 30.3%; Pred. No. 0.6;
 Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;
 QY 18 ITPALRAVPELILK-PLAESQFVGKMMMPAGIALITASTFRSMIDKTLFGLIAY 76
 DB 134 IAFALGVALIGSRVPLAKITFL--MALATIDGAILITITLFTNDLSASLVAVAI 191
 QY 77 AITVAHLGGRRTLLSVAGTIVFVGLV 105
 DB 192 AVLAVALNCGART-----GVYILVGV 214
 RESULT 3
 YEFL SCHPO STANDARD; PRT; 628 AA.
 ID YEFL SCHPO STANDARD; PRT; 628 AA.
 AC 013879;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative transporter CIB3.15C.
 GN SPAC1B3.15C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Frazer A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsle K.,
 RA James K., Jones L., Jones M., Leach S., McDonald S., McLean J.,
 RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 RA Wellens I., Vancraets G., Rieger M., Schaefer M., Mueller-Huer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
 RA Goiffreau A., Cadet E., Diano S., Gloux S., Leclaire V., Moutier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forzberg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the allantoate permease family.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL: Z98598; CAB1241.1; -
 CC PIR: T38034; T38034.

DR GenedB Spombe; SPAC1B3.15C; -
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 157 177 POTENTIAL.
 FT TRANSMEM 203 223 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 454 474 POTENTIAL.
 FT TRANSMEM 483 503 POTENTIAL.
 FT TRANSMEM 511 531 POTENTIAL.
 FT TRANSMEM 542 562 POTENTIAL.
 FT TRANSMEM 583 603 POTENTIAL.
 FT DOMAIN 92 102 POLY-SER.
 FT DOMAIN 385 388 POLY-GLU.
 SQ SEQUENCE 628 AA; 69076 MM; E1040C37445A6D36 CRC64;
 Query Match 15.2%; Score 80; DB 1; Length 628;
 Best Local Similarity 29.2%; Pred. No. 1.3;
 Matches 33; Conservative 22; Mismatches 46; Indels 12; Gaps 5;
 QY 4 DFGCILLVAVCAVITFPAVAPFLIKPLAESQFVGKMMMPA--GILAITASTFRS 61
 DB 448 NFSSTVSVILVAPITWPF--DAIATLVLPD--RFHKRLTFVSGCLFVLAGLITTFVS 504
 QY 62 NA-----IDKTLTFGL--TAAITVAHLGGRRTLLSVAGTIVFVGLVNL 107
 DB 505 NVMGKRYVGLIITLFGIGPTVPIIMTWSSAMGSHSDGVAAAGIAIVGIGNL 557
 RESULT 4
 YEAL ECOLI STANDARD; PRT; 148 AA.
 ID YEAL ECOLI STANDARD; PRT; 148 AA.
 AC P76240; O07965; O07967;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein yeast.
 GN YEAL OR B1789 OR C2194 OR SF7435 OR S1550.
 OS Escherichia coli.
 OS Escherichia coli O6, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NCBI_TaxID=562, 217992, 623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=KL2 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor M., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=KL2;
 RX MEDLINE=97251358; PubMed=8097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Iseno K.,
 RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
 RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
 RA Nishimoto H., Nishio Y., Oshiro T., Saito N., Sempel G., Seki Y.,
 RA Sivasubramam S., Tagami H., Takeda J., Takemoto K., Wada C.,
 RA Yamamoto Y., Horikuchi T.;
 RA "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 RA corresponding to the 40.1-50.0 min region on the linkage map."
 RL DNA Res. 3:379-392(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takuchi M., Tamakoshi A., Tanaka T., Terpsstra P., Tonouchi A.,
 RA Tsubota V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viart A., Wambut R., Wedler H., Wedler H., Weitzenecker T.,
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RT Nature 390:249-256 (1997).
 CC -1- FUNCTION: INVOLVED IN BRANCHED-CHAIN AMINO ACID TRANSPORT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE AZLD/HIT37/HP1330 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; Y11043; CAA71941.1; -
 CC EMBL; U93876; AAB80902.1; -
 CC EMBL; 299117; CAB14611.1; -
 CC PIR; H69592; H69592.
 CC Subtilisin; BG11916; azLD.
 CC InterPro; IPR008407; azLD.
 CC Pfam; PF05437; azLD; 1.
 CC Transport; Amino-acid transport; Transmembrane; Complete proteome.
 FT TRANSMEM 10
 FT TRANSMEM 46
 FT SEQUENCE 110 AA; 11970 MW; 396226FC9D029033 CRC64;
 SQ
 Query Match 18.3%; Score 96.5; DB 1; Length 110;
 Best Local Similarity 27.2%; Pred. No. 0.0059;
 Matches 31; Conservative 28; Mismatches 42; Indels 13; Gaps 6;
 QY 1 MTTFDSCTLLVAVCAVTFALRAVPELIL---KPL-RESQFVGRKMMMPAGIILATLA 56
 DB 1 MTMTMQMTIMAVVLTGLTRFLPMPISGKTPKRYQYGVX---LPSAVIGLVIT 57
 QY 57 STRFSNAIDIKITLFG---LIIVATVVAHLGKRTLSVAGRTVPGVNL 107
 DB 58 YCLKLD--VSLISGSHGIPELVGAADVLAH-LMKRMILSLAGTIVVAVLVL 108
 RESULT 2
 NHAH_ECOLI STANDARD; PRT; 388 AA.
 ID NHAH_ECOLI STANDARD; PRT; 388 AA.
 AC P13738;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NA(+)/H(+)/antporter 1 (sodium/proton antporter 1).
 GN NHAH OR ANT OR B0019.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCB1 TaxID=562;
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86273145; PubMed=2839489;
 RA Karpel R., Olami Y., Taglicht D., Schuldiner S., Padan E.;
 RT "Sequencing of the gene ant which affects the Na⁺/H⁺ antporter
 RT activity in *Escherichia coli*.";
 RL J. Biol. Chem. 263:10408-10414 (1988).
 [2]
 RN REVISIONS.
 RP Padan E.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;

RA MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isoho K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308 (1992).
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perla N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474 (1997).
 [5]
 RN SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=92042080; PubMed=1657980;
 RA Karpel R., Alon T., Glaser G., Schuldiner S., Padan E.;
 RT "Expression of a sodium proton antporter (NHAH) in *Escherichia coli*
 RT is induced by Na⁺ and Li⁺ ions.";
 RL J. Biol. Chem. 266:21753-21759 (1991).
 [6]
 RN FUNCTION, AND PARTIAL SEQUENCE.
 RP MEDLINE=91250446; PubMed=1645730;
 RA Taglicht D., Padan E., Schuldiner S.;
 RT "Overproduction and purification of a functional Na⁺/H⁺ antporter
 RT coded by *nhaA* (ant) from *Escherichia coli*.";
 RL J. Biol. Chem. 266:11289-11294 (1991).
 [7]
 RN MUTAGENESIS OF HISTIDINE RESIDUES.
 RP MEDLINE=93165668; PubMed=8381959;
 RA Gerchman Y., Olami Y., Rimoni A., Taglicht D., Schuldiner S., Padan E.;
 RT "Histidine-226 is part of the pH sensor of NHAH, a Na⁺/H⁺ antporter
 RT in *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1212-1216 (1993).
 CC -1- FUNCTION: THIS IS A NA(+)/H(+) ANTPORTER. IT CAN MEDIANE SODIUM
 CC UPTAKE WHEN A TRANSMEMBRANE PH GRADIENT IS APPLIED. IT PROBABLY
 CC ACTS IN THE REGULATION OF INTERNAL PH AT THE ALKALINE PH RANGE. IT
 CC CATALYZES THE EXCHANGE OF 2H(+) PER NA(+). ITS ACTIVITY IS HIGHLY
 CC DEPENDENT ON THE PH.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- INDUCTION: TRANSCRIPTION STIMULATED BY HIGH NA+ CONCENTRATIONS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; J03879; AAA23448.2; -
 CC EMBL; D10483; BAB96592.1; -
 CC EMBL; AE000112; AAC73130.1; -
 CC EMBL; S67239; AAB20348.1; -
 CC PIR; C64722; C64722.
 CC Ecogene; EGI0652; nhaA.
 CC InterPro; IPR004670; NHAH.
 CC TIGRfam; TIGR00773; NHAH_1.
 CC Transmembrane; Inner membrane; Transport; Antipor; Sodium transport;
 CC Complete proteome.
 FT TRANSMEM 59
 FT TRANSMEM 95
 FT TRANSMEM 125
 FT TRANSMEM 154
 FT TRANSMEM 179
 FT TRANSMEM 199
 FT TRANSMEM 205
 FT TRANSMEM 221
 FT TRANSMEM 254
 FT TRANSMEM 263
 FT TRANSMEM 328
 FT TRANSMEM 360

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:16 ; Search time 6.01671 Seconds
(without alignments)
934.659 Million cell updates/sec

Title: US-10-608-504-5

Sequence: 1 MTTDFSCILLVAVCAVITP.....RTLLSVGAGTIVGVGLNLF 108

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96.5	18.3	110	1	AZLD_BACSU
2	81.5	15.5	388	1	NHAA_BCOLI
3	80	15.2	628	1	YE1F_SCHPO
4	79	15.0	148	1	YE1F_BCOLI
5	75.5	14.4	109	1	YH37_TAEIN
6	75	14.3	1134	1	CV1A_BOVIN
7	74	14.1	171	1	TRPP_CLOAB
8	73.5	14.0	1049	1	ACRB_ECOLI
9	72	13.7	227	1	Y420_NEIMA
10	72	13.7	537	1	QAV_NEUCR
11	70	13.3	342	1	CCR6_CERAE
12	70	13.3	342	1	CCR6_HUMAN
13	70	13.3	342	1	CCR6_PANTR
14	70	13.3	477	1	GTR8_MOUSE
15	70	13.3	652	1	NUSM_PODAN
16	69.5	13.2	521	1	NU2C_SYNY3
17	69.5	13.2	1033	1	ATRA_CANPA
18	69	13.1	342	1	CCR6_MACNE
19	69	13.1	343	1	CCR6_NACPA
20	69	13.1	343	1	CCR6_MACMU
21	69	13.1	462	1	YBXG_BACSU
22	68.5	13.0	507	1	EPOR_MOUSE
23	68	12.9	266	1	UPK_FUSNN
24	68	12.8	714	1	HPPI_AGRTS
25	67.5	12.8	306	1	Y121_WYCGE
26	67.5	12.8	365	1	Y971_MERTJA
27	67.5	12.8	402	1	OPDE_PSEAB
28	67.5	12.8	430	1	SECY_STRAU
29	67	12.7	423	1	YHUV_ECOLI
30	67	12.7	715	1	NUSM_NEUCR
31	66.5	12.6	1032	1	ATRA_MOUSE
32	66.5	12.6	1032	1	ATRA_RAT
33	66.5	12.6	1034	1	ATRA_RABIT

34	66	12.5	510	1	D7A1_MOUSE
35	65.5	12.5	199	1	YPHA_BACSU
36	65.5	12.5	250	1	COBS_SYNP7
37	65.5	12.5	277	1	FMF4_ECOLI
38	65.5	12.5	1033	1	ATRA_PIG
39	65.5	12.5	1034	1	ATRA_HUMAN
40	65	12.4	333	1	Y4M0_RHISN
41	65	12.4	414	1	YGJU_ECOLI
42	64.5	12.3	147	1	RLIS_SYNY3
43	64.5	12.3	430	1	SECY_STACA
44	64.5	12.3	457	1	YERO_ECOLI
45	64.5	12.3	534	1	HUP3_CHIKE

ALIGNMENTS

RESULT 1
AZLD_BACSU STANDARD; PRT: 110 AA.
ID AZLD_BACSU
AC 007923;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Branched-chain amino acid transport protein azld.
GN AZLD OR BS026700.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxId=1423;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BESC1A1;
RX MEDLINE=97431495; PubMed=9267000;
RA Belitsky B.R., Gueterson M.C.U., Sonenshein A.L., von Wachenfeldt C.;
RT "An lrp-like gene of Bacillus subtilis involved in branched-chain
RT amino acid transport.";
RL J. Bacteriol. 179:5448-5457(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97453479; PubMed=9308178;
RA Sorokin A., Bolotin A., Fumelle B., Hilbert H., Lauber J.,
RA Duesterhoeft A., Ehrlich S.D.;
RT "Sequence of the Bacillus subtilis genome region in the vicinity of
RT the lrp operon reveals two new extracytoplasmic function RNA
RT polymerase sigma factors Sigv and Sigz.";
RL Microbiology 143:2939-2943(1997).
[3]
RP SEQUENCE FROM N.A.
RC MEDLINE=98044033; PubMed=9384377;
RX STRAIN=168;
RA Kunst F., Ogawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.C., Bessieres P., Bolotin A., Borcher S.,
RA Bottaris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.O., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Goughly E.J., Grand G.,
RA Guisepi G., Guy B.J., Haga K., Hatach J., Harwood C.R., Henaut A.,
RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasaahara Y., Kleier-Blanchard W., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
RA Parro V., Pohl J.M., Portelle D., Repoport G., Rey W., Reynolds S.,
RA Preecan E., Pujic P., Fumelle B., Rapoport G., Rey W., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadle Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Serron S.J., Serron P., Shin B.S., Solido B.,

submitted to the EMBL Data Library, July 1999

A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: C75003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <KAM>
A:Cross-references: GB:AJ248288; GB:AL096836; NID:95458960; PIDN:CAB50465.1; PID:el51636
A:Experimental source: Strain Orsay
C:Genetics:
A:Gene: PAB1029

Query Match 15.0%; Score 79; DB 2; Length 314;
Best Local Similarity 26.7%; Pred. No. 1.4;
Matches 20; Conservative 17; Mismatches 26; Indels 12; Gaps 1;

QY 4 DFGCILLVAVCAVITFALRAVPFLIKPLRESQFVGKAMMPAGILAITASTFRSNA 63
DB 238 DIIDLMVIGIASIT-----RPIKVDPSIVKVTPIVLMVALITVSLFRNNK 285

QY 64 IDLKITFTGLAVAI 78
DB 286 VGRKTAATLLVYST 300

RESULT 14

AB1068
probable membrane protein STY4872 [imported] - *Salmonella enterica* subsp. *enterica* serov. *Paratyphi*
C:Species: *Salmonella enterica* subsp. *enterica* serovar *typhi*
A>Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB1068
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Houlle, S.; Gao, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov. *Paratyphi*
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB1068
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03361.1; PID:G16505633; GSPDB:GN00176
C:Genetics:
A:Gene: STY4872

Query Match 15.0%; Score 79; DB 2; Length 394;
Best Local Similarity 26.2%; Pred. No. 1.7;
Matches 27; Conservative 25; Mismatches 43; Indels 8; Gaps 4;

QY 9 LLVAVCAVITFALRAVPFLIKPLRESQFVGKAMMPAGILAITASTFRSNAIDLK 67
DB 261 LLIGACVTLSTASDSLLILISIGFATFMTSLVMP--LARQLSAPGNITLGLV 337

QY 68 TITFGLIAVITVAHLG-GRRITLSV---GAGTIVFVGLVN 106
DB 338 TLTGIGQITGLPLASLGSNGASAITNATLCAALFPALIS 380

RESULT 15

A99631
Na+/H antiporter [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 050952)

C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: A99631
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gaevari, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Ref. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genome
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: A99631
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-388 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833440.1; PID:G13359473; GSPDB:GN00154
A:Experimental source: Strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: Ecs0017
C:Superfamily: Na+/H+-exchanging protein nhaA

Query Match 14.9%; Score 78.5; DB 2; Length 388;
Best Local Similarity 30.3%; Pred. No. 1.9;
Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;

QY 18 ITFALRAVPFLIK-PLRESQFVGKAMMPAGILAITASTFRSNAIDLKITFTGLAV 76
DB 134 IAFNLGVALLGSRVPLAKIFL--MALAIDDLGAILIIFVTNDLSMASLGVAAVAI 191
QY 77 AITVAHLGGRITLSVAGTIVFVGLV 105
DB 192 AVLAVALNLGCVRT-----GVYILNGVV 214

Search completed: April 16, 2004, 06:54:26
Job time: 10.4234 secs


```

Query Match      15.5% Score 81.5; DB 2; Length 388;
Best Local Similarity 30.3%; Pred. No. 0.96;
Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;

QY 18 TFPALRAVFLILK-PLRESQFVGKAMMPGAILALTASTFRSNALDITKLTGGLAV 76
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 134 IFALGVALTLDSRPVLAIRKPL--MALAIIDDLGAILITLFTYNDLSMSLVGAVAAT 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 77 AITVAHLIGGRRTLLSVGAGTIVEGLV 105
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 AVLAVALNLGARRT-----GVYILGVV 214
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
C95963
conserved hypothetical membrane protein SMD21531 (imported) - Sinorhizobium meliloti (S)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence revision 24-Aug-2001 #text change 30-Sep-2001

```

```

QY      8  ILVAVCAVTEFLKRAPEFLIKPLRESQFVGWMMMPAGLITAITAFTFRS----- 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      231 VTMLAPVLIPIRSIVLRSP-----QEGASIGKRAPLVEGFLVAFVILAFNSAGLVPV 263

QY      62 -----NAIDKITFTGLIAIVATVVAHLLGRRITLISVAGATIVFGLVNLF 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      284 LASEVMAISRMWALLAGIVAAGKMT-----SLRRVLEVGDDAVALVAETLF 330

RESULT 8
P82726
AAC transporter ATP-binding protein XF1081 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82726
C:RanXm0005: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
      406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: F82726
A>Status: preliminary
A:Molecule type: DNA
A:Access: 1-589 
A:Cross-references: GB:AE003944; GB:AE003849; NID:G9106023; PIDN:AA83891.1; GSPDB:GN000
A:Experimental source: strain 9a5c
R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
B:Simson, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carret,
A:Neto, E.; Docena, C.; El-Dorry, H.; Facinanti, A.P.; Ferreira, A.J.S.
Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohn
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmietti, D.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawada
A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tesharo, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1081

Query Match      15.2%;      Score 80.      DB 2,      Length 589;
Best Local Similarity      34.7%;      Pred. No. 2;

```

Db 58 YCLMD-VSLSGSHGIEPLVGAARVVLH-LMKKNMLLSAGGVVYVMVVL 108

RESULT 2

D86826
Hypothetical protein yqfC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86826
R/Bolotin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weisenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s8
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: D86826
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <STO>
A/Cross-references: GB:AE005176; PID:g12724619; PID:AAK05710.1; GSPDB:GN00146
A/Experimental source: strain IL1403
A/Genetics: yqfC

Query Match 18.0%; Score 94.5; DB 2; Length 108;
Best Local Similarity 29.9%; Pred. No. 0.015;
Matches 32; Conservative 24; Mismatches 42; Indels 9; Gaps 4;

Qy 3 TDFSCILVAVCAVITFALRAVPFLIKPLRESQFVGKAMMPAGILAITAST-FRS 61
Db 2 SSEFSLITIGALITWISRVLPFLTKKSLPQVYVLSFVPIVMSALMISNLFQ 61

Qy 62 NAIDKLTTF-GLIAVAVTVAHLGGRRTLSVGAGTIVFGVNL 107
Db 62 HLGHPVSNMNLASIPVLAAIL--TKNL-----VIVLVGVSL 101

RESULT 3

AH1617
Hypothetical protein lml1481 [imported] - Listeria innocua (strain Clijp11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AH1617
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matcournam, A.; Me
ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AH1617
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <GIA>
A/Cross-references: GB:AL592022; PIDN:CA096712.1; PID:g16413954; GSPDB:GN00178
A/Experimental source: strain Clijp11262
A/Genetics:

Query Match 16.1%; Score 84.5; DB 2; Length 108;
Best Local Similarity 24.8%; Pred. No. 0.15; Mismatches 43; Indels 9; Gaps 3;
Matches 25; Conservative 24; Mismatches 43; Indels 9; Gaps 3;

Qy 9 LTVAVCAVITFALRAVPFLIKPLRESQFVGKAMMPAGILAITASTFRSNAIDK 67
Db 8 LTVIGGLVTFPRVLPFIFVRKQLQBPVIRYLSVPLCIITLALFVQSLITNENSP 67

Qy 68 TLTF-GLIAVAVTVAHLGGRRTLSVGAGTIVFGVNL 107
Db 68 SINENLALPLTIT-----AILTKNLMWIVIGIISM 101

RESULT 4

D87534
Hypothetical protein CC2301 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: D87534
R/Nernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kojic
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapito, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87534
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <STO>
A/Cross-references: GB:AE005673; NID:g13423820; PIDN:AAK24272.1; GSPDB:GN00148
A/Genetics: CC2301

Query Match 15.9%; Score 83.5; DB 2; Length 407;
Best Local Similarity 30.0%; Pred. No. 0.63; Mismatches 38; Indels 19; Gaps 5;
Matches 33; Conservative 20; Mismatches 38; Indels 19; Gaps 5;

Qy 8 ILTVAVCAVITFA---LRAVPFLI-----KPLRESQFVGKAM-MMPAGIL-----AI 53
Db 130 MLVGAVALGLITFADGVLKAIYQVNLNLEIRPDLAMVVKVSMATYPAILLFMCVRI 189

Qy 54 LTASTFRSNAIDKLTTFGLIAVAVTVAHLGGRRTLSVGAGTIVFG 103
Db 190 LDAMNFRG-----RNIALIAVCIYAGATLGADAPIALIIGVAMIG 234

RESULT 5

AC1255
Hypothetical protein lmo1443 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AC1255
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matcournam, A.; Me
ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AC1255
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <GIA>
A/Cross-references: GB:NC_003210; PIDN:CA099521.1; PID:g16410872; GSPDB:GN00177
A/Experimental source: strain EGD-e
A/Genetics: lmo1443

Query Match 15.8%; Score 83; DB 2; Length 108;
Best Local Similarity 26.9%; Pred. No. 0.21; Mismatches 46; Indels 10; Gaps 3;
Matches 29; Conservative 23; Mismatches 46; Indels 10; Gaps 3;

Qy 6 SCILVAVCAVITFALRAVPFLIKPLRESQFVGKAMMPAGILAITASTFRSNAID 65
Db 5 SYTLFVIGGLVTFIRVLPFIFVRKQLQBPVIRYLSVPLCIITLALFVQSLIT 61

Qy 66 LKTLTFG-----LIAVAVTVAHLGGRRTLSVGAGTIVFGVNL 108
Db 62 RNSPFGINENLNASLPTITAIL--TKOLMWIVVIGIISMALIRLP 107

RESULT 6

CE4722
Na+/H+-exchanging protein nha - Escherichia coli (strain K-12)
C/Species: Escherichia coli

1	96.5	18.3	110	1	H65952	branched-chain am
2	94.5	18.0	108	2	D86826	hypothetical prot
3	84.5	16.1	108	2	AH1617	hypothetical prot
4	83.5	15.9	407	2	D87554	hypothetical prot
5	83	15.8	108	2	AC1255	hypothetical prot
6	81.5	15.5	388	2	C64722	Na ⁺ /H ⁺ -exchang
7	80.5	15.3	345	2	C95963	conserved hypoche
8	80	15.2	589	2	F82726	ABC transporter
9	80	15.2	628	2	T38034	probable transpor
10	79	15.0	148	2	E64939	hypothetical prot
11	79	15.0	148	2	B90941	hypothetical prot
12	79	15.0	148	2	B95789	hypothetical prot
13	79	15.0	314	2	C75003	na ⁺ /ca ²⁺ exchang
14	79	15.0	394	2	AB1068	probable membrane
15	78.5	14.9	388	2	AA9631	Na ⁺ /H antiporter
16	78.5	14.9	388	2	A85482	Na ⁺ /H antiporter
17	77.5	14.7	235	2	D8364	cytochrome-c oxid
18	76	14.4	102	2	T35134	hypothetical prot
19	76	14.4	656	2	B20056	glutathione-regul
20	75.5	14.4	109	1	B6041	branched-chain am
21	75.5	14.4	182	2	F72623	hypothetical prot
22	75	14.3	923	2	AC2705	conserved hypoche
23	75	14.3	923	2	C97487	hypothetical prot
24	75	14.3	1134	2	AA1350	adenylate cyclase
25	74.5	14.2	379	2	S67856	Gung protein - Xat
26	74	14.1	171	2	A97344	uncharacterized m
27	74	14.1	324	2	A83940	sugar ABC transpo
28	74	14.1	325	2	D71159	probable Na ⁺ /Ca ²⁺
29	73.5	14.0	443	2	H96939	ATpase related to

30	73.5	14.0	1049	2	B36938	acridi flavin resista
31	73.5	14.0	C90693			acridine efflux pu
32	73.5		1049	2	G85543	acridine efflux pu
33	73	13.9	405	2	A12534	chromate transport
34	73	13.9	409	2	F98347	probable transport
35	72.5	13.8	156	2	B84327	hypothetical prote
36	72.5	13.8	461	2	E81862	probable integral
37	72.5	13.8	461	2	B81080	drug resistance tr
38	72	13.7	227	2	G81014	conserved hypothet
39	72	13.7	537	2	G31277	quinate transport
40	71.5	13.6	101	2	AF2830	conserved hypothet
41	71.5	13.6	104	2	C97608	hypothetical prote
42	71.5	13.6	286	2	C71059	hypothetical prote
43	71.5	13.6	297	2	A83049	hypothetical prote
44	71.5	13.6	351	2	T44798	iron transport mem
45	71.5	13.6	395	2	A82827	conserved hypothet

ALIGNMENTS

RESULT 1

H69592
branched-chain amino acid transport protein azid - *Bacillus subtilis*

C:\species: Bacillus subtilis

C:\Accession: H69592: T44778
C:\Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #seq_change 20-Jul-2000

R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, J.L.; Capriano, V.; Carter, N.M.; Chhabra, R.

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. *Nature* 200 248-256 1997

Nature 350, 247-250, 1991.
A: Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Furuta, S.; Galizzi, A.; Galletti, T.; Hammond, C.B.; Vennart, A.; Wilbert, H.; Holtschmidt, S.; Hosono, S.; Huillo, M.F.

[illegible]

Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Alanchors; Jawerl, U.; Ladačević, V.; Lee, S.W.; Levine, R.; Liu, Y.; Maubou,

Placens M., Duvall C., Poche E., Poce M., Sadate Y., Satō, T.; Scanlon

A Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero, K.; Krieger, M.; Rivola, C.; Roue, B.; Rose, H.; Baader, A.; Bello, L.; Camarero

Authors: M. Tamashiro; A. Tanaka; T. Teranishi; P. Tononi; A. Tosato; V. Uchiyama;

T.; Winters, P.; Wipst, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Yubane, Yoshitaka; Yumoto, F.; Yoshikawa, H.; Danchin A.

A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Accession number: 106890. **URL:** 0804033. **DAED:** 0204377

A;Accession: H69592
A;Reference number: A09380; MWID: 98049033; EMDL: 5004377

A: Molecule type: DNA
A: Stratus: nucleic acid sequence not shown; translation not shown

A:Residues: 1-110 <NON>
A:Cross-references: GB:T299117; GB:AL009126; PIDN:CAB14611.1; PID:g2635115

A: Experimental source: Strain 108
R: Belitsky, B.R.; Gustafson, M.C.U.; Sonenshein, A.L.; von Wachenfeldt, C.

A:Title: An lrp-like gene of *Bacillus subtilis* involved in branched-chain amino acid transamination
J. Bacteriol. 179, 5448-5457, 1997

A/Reference number: Z22837; MOID:97431495; PMID:9287000
A/Accession: T44778

A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-110 <BB>
A/Cross-references: EMBL:Y11043; NID:g1926275; PIDN:CAA71941.1; PID:g1926282

A;Experimental source: strain 1A1
C;Genetics:

A/Gene: azlD
C/Superfamily: branched-chain amino acid transport protein, AzlD type

C/keywords: branched-chain amino acid transport; transmembrane protein

Query Match	18.3%;	Score 96.5;	DB 1;	Length 110;
Best Local Similarity	27.2%;	Pred. No. 0.0095;		

Matches	31;	Conservative	28;	Mismatches	42;	Indels	13;	Gaps	6;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

```

1 MTTDFSCILLVAVCAVITFALRAVPFLL---KPL-RESQFVGKAMMMAGILALITTA 56
  || : : : : | : || : | : : || : : : :

```

Db 1 MNTMTQOMITIAVVLTGIMLTRFLPPIPPSGKPTPKVQYLGA---LPSAVIGLLVI 57

57 STFRSNAIDKLTIFG--LIAVITVAHLLGRRTLISVGAGTIVFVGLNL 107

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rapn

Page 6

[illegible]

Search completed: April 16, 2004, 07:03:20
Job time : 9.52089 secs

Qy 92 LSVGAGTIVF-----VGLVNL 108
Db 163 IIAIGIGIIFGFGNGDAIGLSNLM 187

RESULT 12

US-10-767-701-40503
Sequence 40503, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT FILING DATE: 2004-01-29
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SBO ID NO 40503
LENGTH: 119
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C65105_1.pep
US-10-767-701-40503

Query Match 13.0%; Score 68.5; DB 6; Length 119;
Best Local Similarity 30.8%; Pred. No. 0.57; Indels 33; Gaps 6;
Matches 33; Conservative 15; Mismatches 26; Indels 33; Gaps 6;

Qy 8 ILVVAVCAVITFA--LRAVPLILKPLRESQVGRKMMMPAGILALITASTFRSNAID 65
Db 18 ILSVSIIVAVITFSPGMGAIPTLMNSEI-----LPSI-----KSLGGS 56

Qy 66 LKTLFGLIAVITVAHL-----GGRRTLS---VGAGTIVFVGL 104
Db 57 IATLANWLTSPAITMTNMLTWSVG--TFLSYWVSAPFLIVFVL 101

RESULT 13

US-60-556-841-2298
Sequence 2298, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Adad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 2298
LENGTH: 475
TYPE: PRT
ORGANISM: Pseudomonas putida KT2440
US-60-556-841-2298

Query Match 13.0%; Score 68.5; DB 7; Length 475;
Best Local Similarity 25.9%; Pred. No. 2.9; Indels 25; Gaps 7;
Matches 30; Conservative 24; Mismatches 37; Indels 25; Gaps 7;

Qy 15 CAVITPALRAVPLILKPLRESQVGRKMMMPAGILALITASTFRSNAI-----DL 66
Db 131 CA--SLLWSVHVLVKGKKAFAINQVTVAKVPLLFVIAAFAPADLFTDINGL 188

Qy 67 KTLTFLI-----AVATIVV-----AHLGGRRTLS-VGAGTIV-FVGLVNL 107
Db 189 SNPOFSGVLEOVVNMMLVTVFVFIGIGASVYSGRAQRSDVGAIVIGFLGLAL 244

RESULT 14
US-10-779-543-23512
Sequence 23512, Application US/10779543
GENERAL INFORMATION:
APPLICANT: Williams et al

TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
TITLE OF INVENTION: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
FILE REFERENCE: 2300-21302
CURRENT APPLICATION NUMBER: US/10/779,543
CURRENT FILING DATE: 2004-02-12
PRIOR APPLICATION NUMBER: 10/076,555
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 09/217,471
PRIOR FILING DATE: 1998-12-21
PRIOR APPLICATION NUMBER: 60/068,755
PRIOR FILING DATE: 1997-12-23
PRIOR APPLICATION NUMBER: 60/080,664
PRIOR FILING DATE: 1998-04-03
PRIOR APPLICATION NUMBER: 60/105,234
PRIOR FILING DATE: 1998-10-21
PRIOR APPLICATION NUMBER: 09/297,648
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: PCT/US99/01619
PRIOR FILING DATE: 1999-01-28
PRIOR APPLICATION NUMBER: 60/072,910
PRIOR FILING DATE: 1998-01-28
PRIOR APPLICATION NUMBER: 60/075,954
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/080,114
PRIOR FILING DATE: 1998-03-31
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 23767
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23512
LENGTH: 537
TYPE: PRT
ORGANISM: Homo sapiens
US-10-779-543-23512

Query Match 12.8%; Score 68; DB 6; Length 537;
Best Local Similarity 34.9%; Pred. No. 3.8; Indels 8; Gaps 3;
Matches 22; Conservative 13; Mismatches 20; Indels 8; Gaps 3;

Qy 53 ILTASTFRSN-AIDLKTLFGLIAVITVAHL-----GGRRTLSVG---ACTIVFVGL 104
Db 224 VLKSVGRQSNMAASLSTGVVVKVSTIPATLIVHVGSKTFLCGSSVMAASLVTMG 283

Qy 105 VNL 107
Db 284 VNL 286

RESULT 15

US-10-767-701-43433
Sequence 43433, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 43433
LENGTH: 119
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C71332_1.pep
US-10-767-701-43433

Query Match 12.8%; Score 67.5; DB 6; Length 119;
Best Local Similarity 26.5%; Pred. No. 0.74; Indels 27; Gaps 3;
Matches 26; Conservative 15; Mismatches 30; Indels 27; Gaps 3;

Qy 11 VVAVCAVITPALRAVPLILKPLRESQVGRKMMMPAGILALITASTFRSNAIDLKTLT 70

```

; APPLICANT: Wendy S. Halsey
; APPLICANT: Jon Chambers
; APPLICANT: Alison Isobel Muir
; APPLICANT: Philip Graham Szekeres
; APPLICANT: Usman Shabon
; APPLICANT: Derek J. Bergsma
; APPLICANT: Nabil A. Elshourbagy
; APPLICANT: David Michalovich
; APPLICANT: Pamela A. Lane
; APPLICANT: Menelas N. Pangalos
; APPLICANT: Melanie Robbins
; APPLICANT: David Malcolm Duckworth
; APPLICANT: Jeffrey Hill
; APPLICANT: Ping Tsui
; APPLICANT: Pankaj Agarwal
; APPLICANT: Randall Forrest Smith
; APPLICANT: Lisa Vawter
; APPLICANT: Catherine E. Ellis
; APPLICANT: Manmandeeshwar Gattu
; APPLICANT: John W. Quillen, Jr.
; APPLICANT: Erin M. Toland
; APPLICANT: Steven Michael Foord
; APPLICANT: Han Ngoc Trinh
; APPLICANT: Alexander Taylor
; APPLICANT: Henry Sarau
; APPLICANT: Steven Ruben
; APPLICANT: George H. Poste
; APPLICANT: Michel Louis Souchet
; APPLICANT: Philippe Laurent Robert
; APPLICANT: Stephane Clement Krief
; APPLICANT: Bernard Emile Joseph Gout
; APPLICANT: Bye Mahe
; TITLE OF INVENTION: HUMAN G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: GP-70775B-C1
; CURRENT APPLICATION NUMBER: US/10/667,762
; PRIOR FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/10/407,386
; PRIOR FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/10/278,107
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/10/125,749
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US/09/988,922
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatSeq for Windows Version 3.0
; SEQ ID NO 76
; LENGTH: 342
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-667-762-76

Query Match      13.3%; Score 70; DB 6; Length 342;
Best Local Similarity 27.8%; Pred. No. 1.3;
Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;

QY 7 CILVVAVCAVITFALRAV-----PFLIKPLRESQFVGKAMMPAGILAI 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 36 CMVLVVFVCGVNGSLVVISFYHKLQSLDVFVLNPLADLVFVCTLPFVNAVAGIHEW 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 54 LTAFTFRSNAIDLKTLTF---GLIAVAITY 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 96 VFGVMCKSLGIYITINFTYMSILTCITV 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-787-018-9
; Sequence 9, Application US/10787018
; GENERAL INFORMATION:
; APPLICANT: Goelling, Jennifer
; APPLICANT: Dairaghi, Daniel J.
; APPLICANT: Hanley, Michael
```

```

; APPLICANT: Miao, Zhenhua
; APPLICANT: Talbot, Dale
; APPLICANT: Schall, Thomas J.
; APPLICANT: Chemocentryx, Inc.
; TITLE OF INVENTION: Chemokine Receptor
; FILE REFERENCE: 019934-0007210US
; CURRENT APPLICATION NUMBER: US/10/787,018
; CURRENT FILING DATE: 2004-02-24
; PRIOR APPLICATION NUMBER: US/09/721,495
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/159,015
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 60/159,210
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: US 60/172,979
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 60/173,388
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/186,626
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 09/686,019
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 9
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chemokine receptor STRL33
US-10-787-018-9

Query Match      13.3%; Score 70; DB 6; Length 342;
Best Local Similarity 27.8%; Pred. No. 1.3;
Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;

QY 7 CILVVAVCAVITFALRAV-----PFLIKPLRESQFVGKAMMPAGILAI 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 36 CMVLVVFVCGVNGSLVVISFYHKLQSLDVFVLNPLADLVFVCTLPFVNAVAGIHEW 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 54 LTAFTFRSNAIDLKTLTF---GLIAVAITY 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 96 VFGVMCKSLGIYITINFTYMSILTCITV 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-60-556-841-11117
; Sequence 1117, Application US/60556841
; GENERAL INFORMATION:
; APPLICANT: Abad, Marx S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)
; CURRENT APPLICATION NUMBER: US/60/556,841
; CURRENT FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 12463
; SEQ ID NO 11117
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-556-841-11117

Query Match      13.1%; Score 69; DB 7; Length 462;
Best Local Similarity 22.1%; Pred. No. 2.4;
Matches 32; Conservative 24; Mismatches 45; Indels 44; Gaps 7;

QY 8 ILVVAVCAV-ITFALRAV-PFLIKPLRES-----QFVGKMA-----MMMPAG 49
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 43 VLAAVAICGIFIFFRAMGEMLYEPSTGSAFATGQYIHFPVAGYITAMSNMFOWIIVG 102
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 50 ILAITLTAFTFRSNAI-DKTLTFLGLIAVAITVVAHLGGR-----TL 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 103 MSEITAVGSYTRKMFPPDLPAITPGIVAAVITIGAAITLSVSGFGEFFEMFAIKITVITILM 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...326
 SEQUENCE DESCRIPTION: SEQ ID NO: 3665:
 US-10-417-884A-3665

Query Match 13.8%; Score 72.5; DB 6; Length 326;
 Best Local Similarity 27.4%; Pred. No. 0.64;
 Matches 29; Conservative 16; Mismatches 38; Indels 23; Gaps 4;

QY 8 ILVV---AVCAVTFALRAVPFLILKPLEESQFVG-KAMMMAGILAIITASTRSN 62
 DB 50 ILVTRLPRTICLIVGATSSICGLIMQHTONKFSPTAGTWDASRTGILVAMIFLPG 109
 QY 63 AIDKTLTFGLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 108
 DB 110 A-----SLVRSLTAFCPAF-----AGTLFLSLTRLF 137

RESULT 7

US-10-417-884A-5593
 Sequence 5593, Application US/10417884A
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESS: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5593:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...117
 SEQUENCE DESCRIPTION: SEQ ID NO: 5593:
 US-10-417-884A-5593

Query Match 13.7%; Score 72; DB 6; Length 117;
 Best Local Similarity 25.7%; Pred. No. 0.22;
 Matches 27; Conservative 24; Mismatches 50; Indels 4; Gaps 3;

QY 5 FSCILVAVCAVITFALRAVPFLILKPLEESQFVGKAMMMAGILAIITASTRSNA 63
 DB 14 FETTLTITIGCTLATLSVLPVLKFDLPQPLLEYLSFPIYIMSLMFSSTFQNI 73
 QY 64 IDKTLTF-GLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 107
 DB 74 GHLPOINTENALASVPTLAAILS--KSLIVIVLAGILSLIRL 116

RESULT 8

US-09-804-291A-449
 Sequence 449, Application US/09804291A
 GENERAL INFORMATION:
 APPLICANT: ZOZULA, SERGEY
 TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
 FILE REFERENCE: 100337.54287US
 CURRENT APPLICATION NUMBER: US/09/804,291A
 CURRENT FILING DATE: 2001-03-13
 PRIOR APPLICATION NUMBER: 60/188,914
 PRIOR FILING DATE: 2000-03-13
 PRIOR APPLICATION NUMBER: 60/192,033
 PRIOR FILING DATE: 2000-03-24
 PRIOR APPLICATION NUMBER: 60/198,474
 PRIOR FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/199,335
 PRIOR FILING DATE: 2000-04-24
 PRIOR APPLICATION NUMBER: 60/207,702
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/213,849
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 60/226,534
 PRIOR FILING DATE: 2000-08-16
 PRIOR APPLICATION NUMBER: 60/230,732
 PRIOR FILING DATE: 2000-09-07
 PRIOR APPLICATION NUMBER: 60/266,862
 PRIOR FILING DATE: 2001-02-07
 NUMBER OF SEQ ID NOS: 531
 SOFTWARE: PatentIn Ver. 3.2
 SEQ ID NO: 449
 LENGTH: 313
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-804-291A-449

Query Match 13.5%; Score 71; DB 5; Length 313;
 Best Local Similarity 26.4%; Pred. No. 0.91;
 Matches 29; Conservative 18; Mismatches 41; Indels 22; Gaps 3;

QY 8 ILVVAVCAVITFALRAVPFLILKPLEESQFVGK---MAMMMAGILAIITASTRSNAI 64
 DB 106 IVLGISCCILTMAYRYAICPLKYSITLSPRACTANVGSSWLTGITATTHS--- 162
 QY 65 DKTLTFGLIAVAITVAHLL-----GGRRTLLSVGAGTIVFV 102
 DB 163 ----LIFSLPFRSHPIIFHFLCDILPVLRLASAGKRSEISVWTATVFI 208

RESULT 9

US-10-667-762-76
 Sequence 76, Application US/10667762
 GENERAL INFORMATION:
 APPLICANT: Erding Hu
 APPLICANT: Yuan Zhu
 APPLICANT: Ganesh M. Sathe
 APPLICANT: Joyce Yue Mao

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rapn

Page 2

SEQ ID NO 8672
LENGTH: 568
TYPE: PRT
ORGANISM: Zea mays
US-60-556-841-8672

Query Match 13.9%; Score 73; DB 7; Length 568;
Best Local Similarity 31.3%; Pred. No. 1.1;
Matches 36; Conservative 17; Mismatches 44; Indels 18; Gaps 7;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAILPAST 58
DB 290 ACVLLVAVCAVITF---GMVPYTLGSDAPLAE-FAAKGLKFTVTLISIGAVAGLTITL 345

QY 59 FRSNAIDKLTITFGL-IAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 346 LVGLVYQSR-LYFGLGRDGLPSVFAEVHPTHTPVQSOIWWGCVAAVAGLFNV 399

RESULT 3
US-60-556-841-11808
Sequence 11808, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 11808
LENGTH: 568
TYPE: PRT
ORGANISM: Zea mays
US-60-556-841-11808

Query Match 13.9%; Score 73; DB 7; Length 568;
Best Local Similarity 31.3%; Pred. No. 1.1;
Matches 36; Conservative 17; Mismatches 44; Indels 18; Gaps 7;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAILPAST 58
DB 290 ACVLLVAVCAVITF---GMVPYTLGSDAPLAE-FAAKGLKFTVTLISIGAVAGLTITL 345

QY 59 FRSNAIDKLTITFGL-IAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 346 LVGLVYQSR-LYFGLGRDGLPSVFAEVHPTHTPVQSOIWWGCVAAVAGLFNV 399

RESULT 4
US-60-556-841-9476
Sequence 9476, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 9476
LENGTH: 618
TYPE: PRT
ORGANISM: Oryza sativa
US-60-556-841-9476

Query Match 13.9%; Score 73; DB 7; Length 618;
Best Local Similarity 29.1%; Pred. No. 1.2;
Matches 34; Conservative 17; Mismatches 44; Indels 22; Gaps 6;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAIL 54
DB 321 ACVLLVAVCAVITF---GMVPYTLGSDAPLAE-FAAKGLKFTVTLISIGAVAGLTITL 376
QY 55 TASTFRSNAIDKLTITFGLIAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107

DB 377 LVGLVYQSR-LYFGLGRDGLPSVFAEVHPTHTPVQSOIWWGCVAAVAGLFNV 430

RESULT 5
US-60-556-841-11757
Sequence 11757, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 11757
LENGTH: 618
TYPE: PRT
ORGANISM: Oryza sativa
US-60-556-841-11757

Query Match 13.9%; Score 73; DB 7; Length 618;
Best Local Similarity 29.1%; Pred. No. 1.2;
Matches 34; Conservative 17; Mismatches 44; Indels 22; Gaps 6;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAIL 54
DB 321 ACVLLVAVCAVITF---GMVPYTLGSDAPLAE-FAAKGLKFTVTLISIGAVAGLTITL 376

QY 55 TASTFRSNAIDKLTITFGLIAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 377 LVGLVYQSR-LYFGLGRDGLPSVFAEVHPTHTPVQSOIWWGCVAAVAGLFNV 430

RESULT 6
US-10-417-884A-3665
Sequence 3665, Application US/10417884A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/417,884A
FILING DATE: 17-Apr-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/POCKET NUMBER: GTC-012
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 3665:
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rapn

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: April 16, 2004, 06:53:52 ; Search time 7.52089 Seconds
(without alignments)
386.360 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526
Sequence: 1 MTDFSCITLVAVCAVITF.....RTLSVGAGTIVGVGLNLF 108

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193692 seqs, 26905285 residues

Total number of hits satisfying chosen parameters: 193692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*
1: /cgn2_6/ptodata/1/paa/PCIT NEW COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06 NEW COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07 NEW COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08 NEW COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09 NEW COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10 NEW COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60 NEW COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.5	15.7	156	US-10-724-972A-6988	Sequence 6988, Ap
2	73	13.9	568	US-60-556-841-8672	Sequence 8672, Ap
3	73	13.9	568	US-60-556-841-11808	Sequence 11808, A
4	73	13.9	618	US-60-556-841-9476	Sequence 9476, Ap
5	73	13.9	618	US-60-556-841-11757	Sequence 11757, A
6	72.5	13.8	326	US-10-417-884A-3665	Sequence 3665, Ap
7	72	13.7	117	US-10-417-884A-5593	Sequence 5593, Ap
8	71	13.5	313	US-09-804-291A-449	Sequence 449, App
9	70	13.3	342	US-10-667-762-76	Sequence 76, Appl
10	70	13.3	342	US-10-787-018-9	Sequence 9, Appl
11	69	13.1	462	US-60-556-841-11117	Sequence 11117, A
12	68.5	13.0	119	US-10-767-701-40503	Sequence 40503, A
13	68.5	13.0	475	US-60-556-841-2298	Sequence 2298, Ap
14	68	12.9	537	US-10-779-543-23512	Sequence 23512, A
15	67.5	12.8	119	US-10-767-701-43433	Sequence 43433, A
16	67.5	12.8	432	US-60-556-841-6678	Sequence 6678, Ap
17	67.5	12.8	432	US-60-556-841-6915	Sequence 6915, Ap
18	67	12.7	174	US-60-556-841-8566	Sequence 8566, Ap
19	67	12.7	300	US-60-556-841-11087	Sequence 11087, A
20	67	12.7	447	US-60-556-841-3680	Sequence 3680, Ap
21	67	12.7	448	US-60-556-841-10069	Sequence 10069, A
22	66.5	12.6	398	US-10-724-972A-4715	Sequence 4715, Ap
23	65.5	12.5	364	US-10-767-701-46482	Sequence 46482, A
24	65.5	12.5	351	US-10-724-972A-4781	Sequence 4781, Ap
25	65.5	12.5	440	US-10-724-972A-6421	Sequence 6421, Ap
26	65.5	12.5	486	US-60-556-841-4550	Sequence 4550, Ap

27	65	12.4	129	US-10-767-701-54323	Sequence 54323, A
28	65	12.4	465	US-10-417-884A-5620	Sequence 5620, Ap
29	65	12.4	469	US-60-556-841-4458	Sequence 4458, Ap
30	65	12.4	507	US-60-556-841-4463	Sequence 4463, A
31	65	12.4	508	US-60-556-841-12239	Sequence 12239, A
32	65	12.4	509	US-60-556-841-3141	Sequence 3141, A
33	64.5	12.3	181	US-60-556-841-940	Sequence 940, App
34	64.5	12.3	181	US-60-556-841-9000	Sequence 9000, Ap
35	64.5	12.3	423	US-10-724-972A-4088	Sequence 4088, Ap
36	64.5	12.3	605	US-10-724-972A-4854	Sequence 4854, Ap
37	64	12.2	458	US-10-767-701-52684	Sequence 52684, A
38	64	12.2	425	US-60-556-841-6530	Sequence 6530, A
39	63.5	12.1	297	US-10-724-972A-5519	Sequence 5519, A
40	63.5	12.1	311	US-10-767-701-43663	Sequence 43663, A
41	63.5	12.1	348	US-10-417-884A-6950	Sequence 6950, Ap
42	63.5	12.1	502	US-10-724-972A-7473	Sequence 7473, Ap
43	63.5	12.1	507	PCT-US04-09416-2	Sequence 2, Appl
44	63.5	12.1	518	US-10-724-972A-6261	Sequence 6261, Ap
45	63.5	12.1	572	US-60-556-841-7889	Sequence 7889, Ap

ALIGNMENTS

```
RESULT 1
US-10-724-972A-6988
Sequence 6988, Application US/10724972A
GENERAL INFORMATION:
APPLICANT: Douncette-Stamm, Lynn
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: PAT803-16
CURRENT APPLICATION NUMBER: US/10/724, 972A
PRIOR FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: 09/450, 969
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: 09/134, 001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 7544
SEQ ID NO 6988
LENGTH: 156
TYPE: PRT
ORGANISM: S.epidermidis
US-10-724-972A-6988
Query Match 15.7%; Score 82.5; DB 6; Length 156;
Best Local Similarity 25.7%; Pred. No. 0.019;
Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;
Cy 1 MTDFSCITLVAVCAVITFALRAVPLIKPRRSQFVGKMMMPAGIATITASTFR 60
Db 48 MTDLHLVIV-LGVVTLVIVIPIMSRVNLPAIVIKMSFIP---ITFTLII- 102
Cy 61 SNAIDLTLTFG-----LIAVAIVVAHLGGRITLSVGAGTIVFV 102
Db 103 DGVIGQDHAFGYTLNPLVITAIIVFTWIAIFRSLVITIGG---IFV 148
RESULT 2
US-60-556-841-8672
Sequence 8672, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556, 841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
```

PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 5990
 LENGTH: 494
 TYPE: PRT
 ORGANISM: *Proteus mirabilis*
 US-09-543-681A-5990

Query Match 13.6%; Score 71.5; DB 4; Length 494;
 Best Local Similarity 25.4%; Pred. No. 1.2;
 Matches 33; Conservative 25; Mismatches 43; Indels 29; Gaps 6;

QY 8 ILVVAVCAVITFALRAVPLILKPLRESQFVGK-MAMMPAG-----ILAILTASTF- 59
 DB 250 IIAVVLGGLEPYGLRIFEDIRLVWKNMLAKIHLIGLPRAGENLWILHMTASAFI 309
 QY 60 ---RSNAIDKLTLPGL-----IAVAI---TVVAHLGGR-----RTLASVAGT 98
 DB 310 GLMGVPLAAGTLYFQLSLFMLPGIATISGNEILVGLVGAKRFDPAFTGFKSLKAGV 369
 QY 99 IYFVGLVNL 108
 DB 370 LFTIGVVIY 379

RESULT 13
 US-09-328-352-6931
 Sequence 6931, Application US/09328352
 Patent No. 6562958

GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6931
 LENGTH: 169
 TYPE: PRT
 ORGANISM: *Acinetobacter baumannii*
 US-09-328-352-6931

Query Match 13.5%; Score 71; DB 4; Length 169;
 Best Local Similarity 20.4%; Pred. No. 0.35;
 Matches 28; Conservative 25; Mismatches 48; Indels 36; Gaps 4;

QY 8 ILVVAVC-----AVITR-----LRAVPLILKPLRESQFVGKMMMPAGIILAT 55
 DB 30 VLLVLLICGLSONAAVTAAGILVITKTPNPFPIQAGLNLGILHILIGVLTPIA 89
 QY 56 ASTFRSNAIDKLTLP-GLIAVAITVAHLGGR-----RTL 91
 DB 90 SGLTGESLTKSFISFKSLVAIAIGLVAMLGGRGVKMSQPDVAGLIGTVAGVALL 149
 QY 92 LSYGAGTIVFVGLVNL 108
 DB 150 RGVVPGLIAGLSLF 166

RESULT 14
 US-09-489-039A-7704
 Sequence 7704, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 7704
 LENGTH: 188
 TYPE: PRT
 ORGANISM: *Klebsiella pneumoniae*
 US-09-489-039A-7704

Query Match 13.5%; Score 71; DB 4; Length 188;
 Best Local Similarity 19.8%; Pred. No. 0.4;
 Matches 24; Conservative 23; Mismatches 50; Indels 24; Gaps 2;

QY 12 VAVCAVITFALRAVPLILKPLRESQFVGKMMMPAGIILATSTFRSNAIDKLTLP 71
 DB 64 VASILVLIIVRTPLNAFFPMWEKGLTVGIIILITGVAPIASGLTLPSTLIHFNMM 123
 QY 72 -GLIAVAITVAHLGGRRTL-----SYGAGTIVFVGLVNL 107
 DB 124 KSLIAIYAVGVFVSWLGRGVSLMSQPHLVAGLVGTLVGVALLFRGVVPGLIAGLSL 183
 QY 108 F 108
 DB 184 F 184

RESULT 15
 US-09-328-352-6237
 Sequence 6237, Application US/09328352
 Patent No. 6562958

GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6237
 LENGTH: 259
 TYPE: PRT
 ORGANISM: *Acinetobacter baumannii*
 US-09-328-352-6237

Query Match 13.4%; Score 70.5; DB 4; Length 259;
 Best Local Similarity 22.4%; Pred. No. 0.69;
 Matches 38; Conservative 24; Mismatches 37; Indels 71; Gaps 8;

QY 8 ILVVAVCAVIT-TFA-----LRAVPLILKPL-----RESQFVG 40
 DB 78 LIAVIAICAISSPAGACLSLIPQSVLRPFVVMILVIAITLVKKNFGQVHTGKIT 137
 QY 41 KAMMPAGIILAT-----LTASTFRSNAI-----DKLTLP 70
 DB 138 KMLVLAGISLAIIFYDIFPGTGFIFPIFQVDFLHLSLSKIGNFTNLAAIS 197
 QY 71 F-----GLIAVAITVAHLGGRRTLSVGAG--TIYFVGLVNL 107
 DB 198 FLIPTGAILIHLIGLMAAANVLGSIV-GVNTALKYSGSFRIIFLIVSI 246

Search completed: April 16, 2004, 06:56:35
 Job time : 11.2284 secs

APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 13014
 LENGTH: 1049
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-13014

Query Match 13.8%; Score 72.5; DB 4; Length 1049;
 Best Local Similarity 19.4%; Pred. No. 2.2; Mismatches 32; Indels 53; Gaps 3;
 Matches 26; Conservative 23;

QY 5 FSCILIVAVCAVITFALRAVPLILKPLRESQFVGKAMMPAGILAIITASTR--- 60
 DB 877 YALSLIVFCLALAYESWSIPFSVW-----LVPLGVGVALAATFRGLTN 923
 QY 61 -----SNAI-----DKTLTFGLIAVITVAHL 84
 DB 924 DYFQVGLTTTGLSKAMILIVERPAKDMEXGKGLIETLEAVRMRLPILMTSLAFI 983
 QY 85 LGGRRTLLSVAGT 98
 DB 984 LGVMPLVISSGAS 997

RESULT 10
 US-09-107-532A-5593
 Sequence 5593, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESS: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051,571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277
 INFORMATION FOR SEQ ID NO: 5593:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 117 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...117
 SEQUENCE DESCRIPTION: SEQ ID NO: 5593:
 US-09-107-532A-5593

Query Match 13.7%; Score 72; DB 4; Length 117;
 Best Local Similarity 25.7%; Pred. No. 0.16; Indels 4; Gaps 3;
 Matches 27; Conservative 24; Mismatches 50;

QY 5 FSCILIVAVCAVITFALRAVPLILKPLRESQFVGKAMMPAGILAIITASTERSNA 63
 DB 14 FEBFLITLIGCTLATMRSVLKPFDLQPLEVLSFVIVIMSALMFSLSFTONI 73
 QY 64 IDKTLTF-GLIAVITVAHLGRRITLSVAGTIVGVNL 107
 DB 74 GHLPLINENALASVPTLALILS--KSLIVLVAGILSLIRL 116

RESULT 11
 US-09-252-991A-27266
 Sequence 27266, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196,136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 27266
 LENGTH: 303
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-27266

Query Match 13.6%; Score 71.5; DB 4; Length 303;
 Best Local Similarity 27.4%; Pred. No. 0.63; Indels 43; Gaps 7;
 Matches 34; Conservative 15; Mismatches 32;

QY 4 DFSCILIVAVCAVITFALRAVPLILK-----PLRESQFVG-----KXAMW 45
 DB 35 DFPPLF-----CAIRPALALPLVLRGMPAPFGRIILIGVLGVKFGLLFVGAAQ 89
 QY 46 MEAGILAIITASTERSNAIDKTLTFGLIAVITVAHLGRRITLSVAGTIV 100
 DB 90 MPAGLSLVLSQ-----VFITLIALMLGERSPFGRLGSLAAGLV 134
 QY 101 FVGL 104
 DB 135 LIGL 138

RESULT 12
 US-09-543-681A-5990
 Sequence 5990, Application US/09543681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rat

Page 2

CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6109
LENGTH: 405
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6109

Query Match
Best Local Similarity 31.7%; Pred. No. 0.0012;
Matches 33; Conservative 18; Mismatches 30; Indels 23; Gaps 6;

QY 11 VAVCAVITPALRAVPFLIKPLRE-SQVGVKAMMPAGILV-----AIIITSTRSNAT 64
DB 292 IIAAGCFIVTLLAAV-----REISWFIQIYFLW---GIILGATTPVLTALISRTAE 340
QY 65 DKKTIFGLIAVAITVAHLGRRRTLSVAGTIVFVGLVNLF 108
DB 341 SKQGYILGLVQ-SISQFASWVG-----IALGGILVWPGTLTLF 378

RESULT 3
US-09-134-001C-5172
Sequence 5172, Application US/09134001C
Patent No. 6389370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5172
LENGTH: 156
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5172

Query Match
Best Local Similarity 25.7%; Pred. No. 0.012;
Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;
QY 1 MTTFDSCTILVAVCAVITPALRAVPFLIKPLRESQVGVKAMMPAGILAITSTR 60
DB 48 MTTFDLTLVIV-LCGVVTLLVIVIPRIMISRVNLPALIVKWSFIP---ITLFTALIT- 102
QY 61 SNMIDKLTTFG-----LIAVAITVAHLGRRRTLSVAGTIVFV 102
DB 103 DGVIOCHDAFGYTLNPIYIAILVPIYMLAIFRSLTITLIGS---IFV 148

RESULT 4
US-09-540-236-2360
Sequence 2360, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Bretton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2360
LENGTH: 506
TYPE: PRT

ORGANISM: M. catarrhalis
US-09-540-236-2360

Query Match
Best Local Similarity 14.3%; Score 75; DB 4; Length 506;
Matches 28; Conservative 19; Mismatches 46; Indels 28; Gaps 3;

QY 11 VAVCAVITPALRAVP-----LIKPLRESQVGVKAMMP 46
DB 224 VVMMLAAAFKLSAAPHSVSDVYGAPAPLAFIASFASVAMMALARFVGSFTAM 283
QY 47 PAG---IIAIIITSTRSNATIDKTLTFLGLIAVAITVAHLGRRRTLSVAGTIVFV 103
DB 284 PAFETIIVIVISILMGNLAIRONNIKRM-LAYSIAHIGVALTALISIGASLPVS 342
QY 104 L 104
DB 343 M 343

RESULT 5
US-08-726-214-2
Sequence 2, Application US/08726214
Patent No. 6107076
GENERAL INFORMATION:
APPLICANT: Tang, Wei-Jen
APPLICANT: Gilman, Alfred G.
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2;
SEQUENCE CHARACTERISTICS:
LENGTH: 1134 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-726-214-2

Query Match
Best Local Similarity 14.3%; Score 75; DB 3; Length 1134;
Matches 30; Conservative 20; Mismatches 45; Indels 10; Gaps 5;
QY 12 VAVCAVITPALRAVPFLIKPLRESQVGVKAMMPAGILAITSTRSNAT---DLK 67
DB 127 LALPLSLTFALLCCFPALGPGAGAHAGAAVATADQGWQLLIV-IFVSYALLPVRSLL 185
QY 68 TITFGLIAV--IIVVAHLGRRTL--SVGAGTIVFVGLVNLF 108

Tue Apr 20 06:47:21 2004

US-10-608-504-5.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 10.2284 Seconds
(without alignments)
545.110 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526

Sequence: 1 MTTDFSCILLVAVCAVITF.....RTLLSVGAGTIVFGLVNL 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*

1: /cgm2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgm2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgm2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgm2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgm2_6/prodata/2/1aa/6C.COMB.pep:*
6: /cgm2_6/prodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	526	100.0	108	4	US-09-471-803A-5
2	94.5	18.0	405	4	US-09-543-681A-6109
3	82.5	15.7	156	4	US-09-134-001C-5172
4	75	14.3	506	4	US-09-540-236-2360
5	75	14.3	1134	3	US-08-726-214-2
6	73	13.9	1134	4	US-09-245-039-2
7	73	13.9	590	4	US-09-489-039A-7412
8	72.5	13.8	326	4	US-09-107-532A-3665
9	72.5	13.8	1049	4	US-09-489-039A-13014
10	72	13.7	117	4	US-09-107-532A-5593
11	71.5	13.6	303	4	US-09-252-991A-21266
12	71.5	13.6	484	4	US-09-543-681A-5890
13	71	13.5	169	4	US-09-328-352-6931
14	70.5	13.4	188	4	US-09-489-039A-7704
15	70.5	13.4	259	4	US-09-328-352-6237
16	70.5	13.4	514	4	US-09-543-681A-4255
17	70	13.3	342	2	US-08-742-011-2
18	70	13.3	342	2	US-09-375-384B-5
19	70	13.3	342	3	US-09-116-498-2
20	70	13.3	342	3	US-09-116-498-4
21	70	13.3	342	4	US-09-449-437A-2
22	70	13.3	342	4	US-09-517-605-9
23	69.5	13.2	422	4	US-09-489-039A-7295
24	69.5	13.2	424	4	US-09-489-039A-7646
25	69.5	13.1	839	4	US-09-489-039A-12252
26	68.5	13.0	342	3	US-09-116-498-6
27	68.5	13.0	414	4	US-09-328-352-6961

28	67.5	12.8	233	4	US-09-134-000C-6063	Sequence 6063, Ap
29	67	12.7	195	4	US-09-134-000C-5992	Sequence 5992, Ap
30	66.5	12.6	398	4	US-09-134-001C-4353	Sequence 4353, Ap
31	66.5	12.6	421	4	US-09-489-039A-9655	Sequence 9655, Ap
32	66.5	12.6	463	4	US-09-252-991A-29335	Sequence 29335, A
33	66	12.5	259	4	US-09-540-236-2112	Sequence 2112, Ap
34	66	12.5	485	4	US-09-489-039A-9890	Sequence 9890, Ap
35	66	12.5	659	4	US-09-543-681A-6013	Sequence 6013, Ap
36	65.5	12.5	287	4	US-09-252-991A-23091	Sequence 23091, A
37	65.5	12.5	351	4	US-09-134-001C-4387	Sequence 4387, Ap
38	65.5	12.5	440	4	US-09-134-001C-3286	Sequence 3286, Ap
39	65.5	12.5	1053	4	US-09-543-681A-5981	Sequence 5981, Ap
40	65	12.4	333	4	US-09-562-737-10	Sequence 10, Appl
41	65	12.4	404	4	US-09-328-352-6747	Sequence 6747, Ap
42	65	12.4	465	4	US-09-107-532A-5620	Sequence 5620, Ap
43	64.5	12.3	137	4	US-09-489-039A-10320	Sequence 10320, A
44	64.5	12.3	147	4	US-09-732-210-625	Sequence 625, Ap
45	64.5	12.3	309	4	US-09-540-236-3478	Sequence 3478, Ap

ALIGNMENTS

```
RESULT 1
US-09-471-803A-5
Sequence 5, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERNECHT, NICOLE
APPLICANT: SAMM, HERMANN
APPLICANT: EGGELING, LOETAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471, 803A
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 5
LENGTH: 108
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
FEATURE:
OTHER INFORMATION: ATCC14752
US-09-471-803A-5

Query Match      100.0%; Score 526; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTTDFSCILLVAVCAVITPRLAIVPPLIKPIRESQFVGMMAMMPAGIATITASTFR 60
DB      1 MTTDFSCILLVAVCAVITPRLAIVPPLIKPIRESQFVGMMAMMPAGIATITASTFR 60
QY      61 SNAIDKTLTFGLAVAVITVAHLIGRRITLLSVGAGTIVFGLVNL 108
DB      61 SNAIDKTLTFGLAVAVITVAHLIGRRITLLSVGAGTIVFGLVNL 108

RESULT 2
US-09-543-681A-6109
Sequence 6109, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NOCULEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTUS MIRABII
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
```

XX 20-JUN-2001.
PD 18-DEC-2000; 2000EP-00127688.
XX PF
XX 16-DEC-1999; 99JP-00377484.
XX PR 07-APR-2000; 2000JP-00159162.
XX PR 03-AUG-2000; 2000JP-00280988.
XX
XX PA (KYOW) KYOWA HAKKO KOGYO KK.
XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX DR MPI; 2001-376931/40.
XX N-PSDB; AAH67841.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX PS Claim 17; SEQ ID NO 6376; 246bp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium bacterium, measuring expression amount and analyzing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium bacterium, and identifying a homologue of a gene derived from
XX Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described in the
XX exemplification of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the European Patent Office
SQ Sequence 440 AA;

Query Match 14.5%; Score 76.5; DB 4; Length 440;
Best Local Similarity 26.1%; Pred.No. 1.1; Mismatches 7;
Matches 35; Conservative 19; Indels 45; Gaps 7;
QY 11 VVAVCAVITFALRAVPFLIKELRESQ-----FVGKMA-MMP----- 47
DB 202 LVAVLAAMVFLVSAIPALIRVPEIRAQVAEDHPKGLAAVMDLFGQIAELWKQDRNSVY 261
QY 48 -----AGIL---AIIITASTFRSNALIDLTITLFGILAAVAITVAHLGGRITLL 92
DB 262 FLIAAAVPRDGLAGVFTFGAILAVSVYGISAGDV--LLEGVAAVNVSAIGALIGS--FL 316
QY 93 S--VGAGTIVFVGL 104
DB 317 DDRVGPXPIIILSL 330

Search completed: April 16, 2004, 06:53:15
Job time : 34.3844 secs

CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs; or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 464 AA:
 SQ Query Match 14.6%; Score 77; DB 6; Length 464;
 Best Local Similarity 28.3%; Pred. No. 0.99;
 Matches 30; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

QY 17 VTTPFLRAVPFLIK-----PLRE---SQFGKAMMPAGILALIPASTFR 60
 DB 229 LTFPLSLFAFLLLEHSHKPMINLNFKIRFTSALVALAOPFYIGVITL--PTFF 286
 QY 61 SNAIDKLTLPGLIAVAITVAHLGGRTLL--SVGAGTIVFGL 104
 DB 287 TTIGKTELDALILPLMSLVFICGGLSVYINQGLPFLIVFGL 332

RESULT 14
 AAB79004
 ID AAB79004 standard; protein; 440 AA.
 XX
 AC AAB79004;
 XX
 DT 30-APR-2001 (first entry)
 XX
 DE C. glutamicum SRT protein sequence SEQ ID NO:268.
 XX

KW Corynebacterium glutamicum; stress; resistance; tolerance; SRT;
 KW fine chemical production; organic acid; proteogenic amino acid;
 KW nonproteogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW evolutionary study; environmental hazard; fermentation.
 XX
 OS Corynebacterium glutamicum.
 XX

XX
 PN MO200100804-A2.
 XX

XX
 PD 04-JAN-2001.
 XX

XX
 PF 23-JUN-2000; 2000MO-1B000922.
 XX

XX 25-JUN-1999; 99US-0141031P.
 PR 01-JUL-1999; 99DE-0103042P.
 PR 01-JUL-1999; 99US-0142682P.
 PR 08-JUL-1999; 99DE-01031413.
 PR 08-JUL-1999; 99DE-01031457.
 PR 08-JUL-1999; 99DE-01031541.
 PR 09-JUL-1999; 99DE-01032209.
 PR 14-JUL-1999; 99DE-01032914.
 PR 27-AUG-1999; 99DE-01040764.
 PR 31-AUG-1999; 99US-0151214P.
 XX
 XX 99DE-01041382.
 XX
 PA (BADI) BASF AG.
 XX

PI Pompeius M, Kroeger B, Schroeder H, Zelder O, Habermayer G;
 PI Lee H, Kim H;
 XX
 DR WPI; 2001-061972/07.
 DR N-PSDB; AAF71117.
 XX
 PT New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
 PT tolerance or resistance protein, for production or modulation of
 PT production of fine chemicals, such as, e.g. amino acids, lipids,
 PT carbohydrates, or enzymes.
 XX
 PS Claim 20; Page 481-482; 526pp; English.

XX AAF70984 to AAF71113 encode the Corynebacterium glutamicum stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteogenic or nonproteogenic amino acid (preferred),
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polyketide, or an enzyme. The fine
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (II) encoded by them are used for diagnosing the presence or
 CC activity of Corynebacterium diptheriae. (I), (II), (III) and host cells
 CC containing them can be used to map the genomes of organisms related to C.
 CC glutamicum, to identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determination of SRT protein regions required
 CC for function, in modulating the SRT protein activity, and in modulating
 CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
 CC survive in an environment that is normally environmentally or chemically
 CC hazardous to it. (I) and protein molecules encoded by it increase the
 CC survival of C. glutamicum to chemical and environmental hazards and
 CC provide a means for continued growth and multiplication in large scale
 CC fermentative growth conditions. By increasing the growth rate or
 CC maintaining a normal growth rate in poor or toxic conditions, the yield,
 CC production and/or efficiency or production of fine chemicals from a
 CC culture may be increased
 CC
 XX
 SQ Sequence 440 AA;

Query Match 14.5%; Score 76.5; DB 4; Length 440;
 Best Local Similarity 26.1%; Pred. No. 1.1;
 Matches 35; Conservative 19; Mismatches 35; Indels 45; Gaps 7;

QY 11 VVAVCAVITPALRAVPFLIKPLERSQ-----FVGMA-MMP----- 47
 DB 202 LVAVLAAMVFLVSAIPALRVPEIEAQAADHPKGLIAAYKDLFGQIAELMKQDRNSVY 261
 QY 48 -----AAIL---AIIITASTFRSNAIDKLTLPGLIAVAITVAHLGGRTLL 92
 DB 262 FLIAATVFRDGLAGVFPFGAILAVSYGLSGDV--LIFVAAVNVGALCALGG---FL 316

QY 93 S--VGAGTIVFGL 104
 DB 317 DDRVGPRIILISL 330

RESULT 15

AAAG92622
 ID AAAG92622 standard; protein; 440 AA.
 XX

XX
 AC AAAG92622;
 XX

XX
 DT 26-SEP-2001 (first entry)
 XX

XX C glutamicum protein fragment SEQ ID NO: 6376.
 DE

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX
 OS Corynebacterium glutamicum.
 XX

XX
 PN EPI108790-A2.
 XX

RESULT 12
ABM67632
ID ABM67632 standard; protein; 447 AA
XX
AC ABM67632;

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis
 detection; food; gene expression; plant; animal; microorganism; toxin
 antibiotic; biopesticide; virulence factor; disease model; plague;
 whooping cough.

WO200294867-A2.

28-NOV-2002

07-FEB-2002; 2002WO-IB003040.

07-FEB-2001; 2001FR-00001659.

(INSP) INST PASTEUR.
(CNRS) CNRS CENT NAT RECH SCI.

Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A, Buchrieser C;

WPI; 2003-148459/14.

Genomic sequence of *Photornabodus luminescens* and encoded polypeptides useful e.g. as therapeutic antimicrobials and agricultural pesticides

Claim 2; SEQ ID NO 729; 1205bp; French.

The *inv* gene relates to the isolation of genes and their encoded proteins from *Photobacterium luminescens*. The isolated sequences are sources of probes and primers for detecting the genome of *P. luminescens* and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than *P. luminescens* and are able to alter response or sensitivity to toxins and antibiotics produced by *P. luminescens*. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibiotics useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which *P. luminescens* is a model (particularly plague and whooping cough). This sequence represents one of the isolated *P. luminescens* proteins

SQ Sequence 447 AA;
 Query Match 14.7%; Score 77.5; DB 6; Length 447;
 Best Local Similarity 24.8%; Pred. No. 0.82;
 Matches 34; Conservative 25; Mismatches 47; Indels 31; Gaps 5.

QY 2 TTBSCILVAVCAVTFAALRAVEFLIKPLRESQ-----FVGKMMAMMPAG 49

:| | : | | | : | | :

Db 196 SIVFGRI VAVLLFCLLSYGLR-KFIIAGFLRWSQGMGKILYIGLPAAGBNVVMILQY 254

```

50 ILALITASTFRSNAIDKLTLEGL-----LAVAI--IVVAHLLDGR-----RIL 91

```

DB 255 MAIAFIGNGEISLAQTLYFQLSLFVLMFGIAISIGNEIMVGHLVGAKRFEDAYIRGL 314

	92	LSVGAGTIVEGLVNLF	108
QY			

Db 315 NSLKIGCIVTIGVTA 331

РЕСПУБЛИКА

ABU44538

XX
AC
APR 14 1968

XX
XX
F
C
C
C
(
4
4
4
)

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

XX

2000

33

XXXXXXXXXXXX

X

PR 06-SEP-2001; 2001US-00948993

PR 08-FEB-2002; 2002US-00072851

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

2 X

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH; ...

DR WPI: 2003-029926/02.

XX

PT for homologous nucleic acids required for cellular proliferation to

2
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
8

the 6213 antisense sequences given in the specification where expression

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC proliferation or the activity of a gene in an operon required for

CC the gene product or that has an activity against a biological pathway

CC identifying a gene required for cellular proliferation or the biological

XX PS Claim 6; SEQ ID NO 665; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC monocytes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC monocytes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC monocytes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 108 AA;

Query Match 15.8%; Score 83; DB 5; Length 108;
 Best Local Similarity 26.9%; Pred. No. 0.031;
 Matches 29; Conservative 23; Mismatches 46; Indels 10; Gaps 3;

QY 6 SCILVAVCAVTFPLRAVPLIKPLRESQVGMAMMAGLALITASTFRSND 65
 DB 5 SYTLFVIGGLVTFPRVPLFVRKQLDPVIVRYSVP--LCITLALFVOSLIT 61
 QY 66 LKTLTFG-----LIAVAIVVAHLGSRRLTSLVGAGTIVFGLVNF 108
 DB 62 RENSFGINENMLASLPITITAIL--TKNIMWIVVGIISMLRLF 107

RESULT 10
 ABP40327
 ID ABP40327 standard; protein; 156 AA.

XX AC ABP40327;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5172.
 XX
 KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 XX
 KM antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-00134001.
 XX
 PR 14-AUG-1997; 97US-0055779P.
 PR 08-NOV-1997; 97US-0064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 PI Doucette-Stamm LA, Bush D;
 DR WPI, 2002-381255/41.
 DR N-PSDB; ABN92872.
 XX
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermis
 XX polypeptide, useful for diagnosing and treating bacterial infections.
 PS Claim 1; SEQ ID NO 5172; 267pp; English.
 XX

CC AAB90538 to AAB93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in AAB935124 to AAB97960. The *S. epidermidis* sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,
 CC particularly *S. epidermidis* infections. The sequences can be used to
 CC screen for compounds able to interfere with the *S. epidermidis* life cycle
 CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site
 CC
 SQ Sequence 156 AA;

Query Match 15.7%; Score 82.5; DB 5; Length 156;
 Best Local Similarity 25.7%; Pred. No. 0.056;
 Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

QY 1 MTDFSCILVAVCAVTFPLRAVPLIKPLRESQVGMAMMAGLALITASTFR 60
 DB 48 MTDDLHLVLIV-LCGVTLVLRVIFPIMISRYNLPAIVIKMISFIP--ITLFTALII- 102
 QY 61 SNAIDLKTTFG-----LIAVAIVVAHLGSRRLTSLVGAGTIVFV 102
 DB 103 DGVICQHDHAFGYTLNLPYIIALVPTVMAIFRSLTVITLGG--IFV 148

RESULT 11
 AAB96384
 ID AAB96384 standard; protein; 314 AA.

XX AC AAB96384;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Putative P. abyssi sodium/calcium antiporter.
 XX
 KW Hyperthermophilic archaeon; hyperthermophilic protein.
 XX
 OS Pyrococcus abyssi.
 XX
 PN FR2792651-A1.
 XX
 PD 27-OCT-2000.
 XX
 PF 21-APR-1999; 99FR-00005034.
 XX
 PR 21-APR-1999; 99FR-00005034.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA (IFREMER) IFREMER INST FR RECH EXPL MER.
 XX
 PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O,
 PI Querrelou J, Weissendach J, Saurin W, Heilig R;
 DR WPI, 2001-126236/14.
 XX
 PT New nucleotide sequences isolated from Pyrococcus abyssi encode proteins
 PT useful in industry.
 XX
 PS Claim 7; Page 1068-1069; 1657pp; French.
 XX
 CC The present invention relates to the genomic sequence of *Pyrococcus*
 CC abyssi (see AAF86431 and AAH41223-7) and *P. abyssi* proteins. *P. abyssi* is
 CC a hyperthermophilic archaeon, which is isolated from deep-sea
 CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.
 CC The proteins of the present invention have various potential industrial
 CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade. Note: This patent is in the same patent family as
 CC WO200065062, which contains additional sequences as shown in AAB99132-
 CC AAB99143, AAH75903-AAH75920 and AAG66436
 XX
 SQ Sequence 314 AA;

XX SQ Sequence 297 AA;
 Query Match 16.6%; Score 87.5; DB 6; Length 297;
 Best Local Similarity 30.3%; Pred. No. 0.029;
 Matches 37; Conservative 21; Mismatches 39; Indels 25; Gaps 8;

QY 7 CILVVAVCAVITPALRAVFLIK-----PLRSQFVG---KMAWMPAGILAILITA 56
 DB 140 CPLL---VCLVLTLLMKTEIGLVNSTGDNIPMSANGVNDTKIVGWSNGLIALCG 196
 QY 57 STFRSN-----AIDLKITLFFGLIAVAIT-VVAH--LLGGRRTLLSVAGTIVF-VGLVN 106
 DB 197 SLFQNDGSDVTSQGTIVGSLSSVILAEVLIHDLITGR--LISIGIAIVRLIILN 254

QY 107 LF 108
 DB 255 IX 256

RESULT 6
 ABR42400
 ID ABR42400 standard; protein; 563 AA.
 XX
 AC ABR42400;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Human type I adenylylate cyclase splice variant.
 XX
 KW Adenylylate cyclase; enzyme; human; neuroprotective; nootropic;
 KW antiParkinsonian; cerebroprotective; analgesic; antidiabetic;
 KW antiinflammatory; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003038093-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 29-OCT-2002; 2002MO-EP012059.
 XX
 PR 29-OCT-2001; 2001US-0330710P.
 PR 23-JUN-2002; 2002US-0350019P.
 PR 19-JUN-2002; 2002US-0389538P.
 PR 26-SEP-2002; 2002US-0413597P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WPI; 2003-441359/41.
 DR N-PSDB; ACCS8112.
 XX
 PT New polynucleotide encoding a type I adenylylate cyclase polypeptide,
 PT useful for diagnosing, preventing or treating diseases associated with
 PT type I adenylylate cyclase dysfunction, e.g. CNS disease or diabetes.
 XX
 PS Claim 1; Page 139-140; 167pp; English.
 XX
 CC The present sequence is the protein sequence of a splice variant of human
 CC type I adenylylate cyclase (see also ABR42339), an enzyme that catalyzes
 CC the synthesis of cAMP from ATP. The invention provides reagents that
 CC regulate human type I adenylylate cyclase and reagents which bind human
 CC type I adenylylate cyclase gene products. These reagents, and vectors
 CC containing polynucleotides encoding e.g. present sequence, can be used to
 CC modulate the activity of human type I adenylylate cyclase in a disease,
 CC especially a central nervous system (CNS) disorder, diabetes and chronic
 CC obstructive pulmonary disorder (claimed). The CNS disorders may include
 CC Alzheimer's disease, Parkinson's disease, stroke or neuropathic pain.
 CC Type I adenylylate cyclase can be obtained by recombinant methods, and used
 CC to screen for activator and inhibitor compounds

XX SQ Sequence 563 AA;
 Query Match 16.2%; Score 85; DB 6; Length 563;
 Best Local Similarity 31.4%; Pred. No. 0.13;
 Matches 33; Conservative 18; Mismatches 44; Indels 10; Gaps 5;

QY 12 VAVCAVITPALRAVFLIKPLRSQFVGKMAWMPAGILAILITASTFRSNAI---DLK 67
 DB 125 LALLFSLTFLALCCPPLAGGPARGSAGAGPATAEQGWOLLV-TFVSVALLPYRSIL 183
 QY 68 TLFFGLIAVA--ITVVAHLGGRITL--SVAGTIVFGLVNL 108
 DB 184 AIGGLVAASHLVLVTRTLVPARPLRLTGLNALLFVG-VNMY 227

RESULT 7
 AAE17133
 ID AAE17133 standard; protein; 1119 AA.
 XX
 AC AAE17133;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Human adenylyl and guanylyl cyclase (ADGUC)-5.
 XX
 KW Human; adenylyl and guanylyl cyclase; ADGUC-5; cardiovascular disorder;
 KW angina pectoris; myocardial infarction; vision disorder; keratitis;
 KW iritis; cataract; neurological disorder; epilepsy; Alzheimer's disease;
 KW Pick's disease; stroke; mental disorder; mood and anxiety disorder;
 KW reproductive disorder; infertility; endometriosis; impotence; asthma;
 KW smooth muscle disorder; migraine; bacterial infection; gene therapy;
 KW transgenic animal; vaccine; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN WO200202757-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 26-JUN-2001; 2001MO-US020491.
 XX
 PR 29-JUN-2000; 2000US-0215476P.
 PR 04-AUG-2000; 2000US-0223545P.
 PR 31-AUG-2000; 2000US-0229876P.
 PR 22-SEP-2000; 2000US-0234838P.
 PR 29-SEP-2000; 2000US-0236483P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Ganahi AR, Tribouley C, Ding L, Lu DM, Lee EA, Yue H, Yang J,
 PI Baughn MR, Thornton M, Yao MG, Walia NK, Tang YT, Elliott VS, Lu Y;
 XX
 DR WPI; 2002-154740/20.
 DR N-PSDB; AAD28061.
 XX
 PT Novel human adenylyl and guanylyl cyclases and polynucleotides encoding
 PT the cyclases, useful for treating, diagnosing or preventing
 PT cardiovascular, neurological, vision, reproduction and smooth muscle
 PT disorders.

Query Match	24.7%	Score 130;	DB 4;	Length 115;
Best Local Similarity	35.0%	Pred. No. 5.1e-08;		
Matches	36;	Conservative 16;	Mismatches 47;	Indels 4;
				Gaps 2;

Qy 8 ILLVNA- ---CAVTFLALRAPELLEPLEBESQFVGGAAMMMAGIALIATSTPFSNAI 64
| | | | |
Db 12 LNVAAVALPIAIIITTLRLRFFPAAMKRVANSNQMLGRTMPEGVWVVIVITLEGQS 71
| | | | |
Qy 65 DLKTLTFGLIAVAITTVVAHLGGRRITLSVAGTIYFVGLVNL 107
| | | | |
Db 72 APGGVGSILIAVAFITALIHWLKSGAGSIYG -GTLYMYELVN 113
| | | | |

RESULT 4

ID ABB54958 standard; protein; 108 AA.

AC ABB54958;

DT	29-AUG-2003	(revised)
DT	16-MAY-2002	(first entry)

DE Lactococcus lactis protein yqfC.

Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

OS Lactococcus lactis; IL1403.

AA
PN FR2807446-A1.

AA
PD 12-OCT-2001

11-APR-2000: 2000FR-00004630.

XX
PR 11-APR-2000: 2000FR-00004630

XX
PA (TNRG) INRA INST NAT RECH AGRONOMIQUE.

XX	Sorokine A.	Renault P.	Ehrlich SD.
PI	Bojorine A.		

XX
DB
WPT: 2002-043418/06

AA New nucleotide sequence useful in the identification or *Lactococcus*
PT
PT facts and related species.

XX
PS C]ajm 6: SEO ID NO 1660: 2504pb: French.

The present invention is related to a *Lactococcus lactis* nucleotide sequence (AB490521) and related proteins (AB55300-AB55521). The nucleic acid sequence is useful in the detection and/or amplification of nucleic acid sequence, particularly to identify *Lactococcus lactis* or related species. The proteins of the invention are useful for the biosynthesis or biodegradation of a composition of interest. The invention helps research in lactic bacteria, particularly useful in the production of yogurt and cheese. Note: The sequence data for this patent is based on equivalent patent WO20017734 (published as WO/01) which is available in electronic format directly from WIPO at [ftp://wipo.int/pub/published_pat_sequences](http://wipo.int/pub/published_pat_sequences). (Updated on 29-AUG-2003 to standardise OS field)

Sequence 108 AA:

Query Match	18.0%	Score 94.5	DB 5	Length 108
-------------	-------	------------	------	------------

Matches	32;	Conservative	24;	Mismatches	42;	Indels	9;	Gaps	4
3	TDFSCILLVAVCAVITFAIRAPPEILIKPEESQEVGKMMMPAGILAITAST-FRS	61.							

```
D6      2 SFEPISTLTIGCAIVTISKRLPFIILKKKSLDPQIVELRYLSPVIVIMASB
QY      62 NAIDKTLTF-GLAVAITVAHLLGGRTLLSVGAGTIIVEGVNL 107
DB      62 HGHGPSVMNNLASIPFLAII--TKNLL-----YIIVAGFSL 101
```

RESULT 5
ABU45714
ID ABU45714 standard; protein; 297 AA

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #31241.

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

Streptococcus pneumoniae.

AA WO200277183-A2
PN

03-OCT-2002
AA
PD

21-MAR-2002: 2002WO-US009107.

AA 21-MAR-2001: 2001US-00815242.
PR

PR 06-SEP-2001; 2001US-00348993.
PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002TS-0362699P.

XX
PA (EITTB-) EITTB PHARM INC

	Wang T.	Zamudio C	Majone G
XX			
BT			

PI Wang L,	Zamudio C,	Maldonado
PI Wang D,	Trawick JD,	Carlson

XX
DB WPT: 2003-029926/02

DR N-PSDB; ACA49584.
YY

PT New antisense nuc
PT for homologous nu

PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 73638; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WPO at http://wipo.int/pubn/published_pat_sequences

CC especially overexpressed; and (d) method for isolating the brnE and/or
CC brnF genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence represents
CC the Corynebacterium glutamicum ATCC 14752 brnE protein described in the
CC method of the invention
CC
XX

Sequence 108 AA;

Query Match 100.0%; Score 526; DB 4; Length 108;

Best Local Similarity 100.0%; Pred. No. 4.1e-57; Mismatches 0; Indels 0; Gaps 0;

Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTDFSCILLVAVCAVITFALRAVPFLIKPRLRSQFVGKAMMPAGILAITASTFR 60
1 MTTDFSCILLVAVCAVITFALRAVPFLIKPRLRSQFVGKAMMPAGILAITASTFR 60
DB 61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108

QY 61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108
61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108
DB

RESULT 2

AAG90036 ID AAG90036 standard; protein; 108 AA.

AC AAG90036;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 3790.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP1108790-A2.

XX 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00155162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

DR N-PSDB; AAH65255.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 3790; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homolog of a gene derived from

CC Coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the European Patent Office

XX Sequence 108 AA;

Query Match 100.0%; Score 526; DB 4; Length 108;

Best Local Similarity 100.0%; Pred. No. 4.1e-57; Mismatches 0; Indels 0; Gaps 0;

Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTDFSCILLVAVCAVITFALRAVPFLIKPRLRSQFVGKAMMPAGILAITASTFR 60
1 MTTDFSCILLVAVCAVITFALRAVPFLIKPRLRSQFVGKAMMPAGILAITASTFR 60
DB 61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108

QY 61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108
61 SNAIDKLTLPGLIAVAITVAHLIGRRITLSVGAGTIVFGVLNLF 108
DB

RESULT 3

AAG93153 ID AAG93153 standard; protein; 115 AA.

AC AAG93153;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6907.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP1108790-A2.

XX 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00155162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

PA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

DR N-PSDB; AAH68372.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 6907; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homolog of a gene derived from

CC Coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC electronic format directly from the European Patent Office

XX Sequence 115 AA;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:12 ; Search time 30.3844 Seconds
(without alignments)
1004.303 Million cell updates/sec

Title: US-10-608-504-5
Perfect score: 526
Sequence: 1 MTPPSCILVAVCAVITF.....RTLSVGAGTIVFVGLNLF 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20048:*
8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	526	100.0	108	4	AA86248 C. glutam
2	526	100.0	108	4	AA86248 C. glutam
3	130	24.7	115	4	AA86248 C. glutam
4	94.5	18.0	108	5	AB854958 Lactococ
5	87.5	16.6	297	6	ABU45714 Protein e
6	85	16.2	563	6	ABR42400 Human typ
7	85	16.2	1119	5	AA871333 Human ade
8	85	16.2	1119	5	ABR42399 Human typ
9	83	15.8	108	5	AB847960 Listeria
10	82.5	15.7	156	5	ABP40327 Staphyloc
11	79	15.0	314	4	AA867632 Photocorb
12	77.5	14.7	447	6	ABU4538 Protein e
13	77	14.6	464	4	AA879004 C. glutam
14	76.5	14.5	440	4	AA862622 C. glutam
15	76.5	14.5	440	4	AA862622 C. glutam
16	75	14.3	415	6	ABU21311 Protein e
17	75	14.3	1134	3	AA862005 Adenylat
18	75	14.3	1134	3	AA862005 Adenylat
19	74.5	14.2	379	1	AA870461 Sequence
20	74	14.1	207	4	AB869858 S. cinna
21	74	14.1	227	6	ABP7220 N. gonorr
22	74	14.1	254	4	AA871711 Human oif
23	74	14.1	311	5	AB851594 Human G P
24	74	14.1	312	4	AA871956 Human oif
25	74	14.1	313	6	ABU11194 Human G-P

26	74	14.1	313	7	ADC85887	Adc85887 Human GPC
27	74	14.1	317	6	ABR01631	ABR01631 Human G P
28	74	14.1	318	4	AA872535	AA872535 Human OR
29	74	14.1	318	4	AA872225	AA872225 Human OLF
30	74	14.1	330	7	ADC86129	ADC86129 Human GPC
31	73.5	14.0	187	6	ADB12425	ADB12425 A11010c
32	73.5	14.0	443	6	ABU24119	ABU24119 Protein e
33	73.5	14.0	1049	4	AA872855	AA872855 E. coli a
34	73.5	14.0	1049	4	AAU34479	AAU34479 E. coli c
35	73.5	14.0	1049	6	ABU15219	ABU15219 Protein e
36	73	13.9	109	6	ABM72816	ABM72816 Staphyloc
37	72.5	13.8	326	7	ADC94038	ADC94038 E. faeciu
38	72.5	13.8	1048	6	ABU32055	ABU32055 Protein e
39	72.5	13.8	1048	6	ABU28297	ABU28297 Protein e
40	72	13.7	117	7	ADC95966	ADC95966 E. faeciu
41	72	13.7	464	4	AB867164	AB867164 Drosophi
42	72	13.7	558	6	ABR42402	ABR42402 Human typ
43	72	13.7	704	6	ABU21547	ABU21547 Protein e
44	72	13.7	1114	6	ABR42401	ABR42401 Human typ
45	71.5	13.6	1049	4	AAU38126	AAU38126 Salmonell

ALIGNMENTS

RESULT 1
ID AAB86248 standard; protein; 108 AA.

XX AC AAB86248;
XX DT 05-SEP-2001 (first entry)
XX DE C. glutamicum brnE protein.
XX DE L-amino acid production; brnF; brnG; branched-chain amino acid;
XX KW Corynebacterium; leucine; isoleucine; valine; medicine;
XX KW animal nutrition.
XX OS Corynebacterium glutamicum.
XX PN EPI096010-A1.
XX PD 02-MAY-2001.
XX PF 11-OCT-2000; 2000EP-00122057.
XX PR 27-OCT-1999; 99DE-01051708.
XX PA (DEGS) DEGUSSA AG.
XX PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX PI Kennertknecht N, Eggeling L, Sahn H, Pfeifferle W,
XX WP: 2001-391595/42.
XX DR N-PSDB; AAH21109, AAH21111.
XX PT New export genes from coryneform bacteria, useful for increasing
XX PT Fermentative production of branched-chain amino acids.
XX PS Claim 1a; Page 17; 23pp; German.

This invention describes a novel isolated polynucleotide (I) containing at least one sequence that (i) is 70% identical with a sequence that encodes at least one of 251 amino acid (aa) (3) or 108 aa (5) polypeptides; (ii) encodes a polypeptide at least 70% identical with (3) or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least 15 consecutive bases from (i)-(iii). The invention also describes (a) protein derived from sequences of 753 bp (2) or 324 bp (4); (b) coryneform microorganisms, especially Corynebacterium, transformed with one or more (II), where these are replicative DNA; (c) production of branched-chain L-aa by fermentation of coryneform bacteria in which the brnE and/or brnF genes (or equivalent sequences) are amplified.

```

Db      555 GATATATACCAAGATGTTTTCATTTGCTTGGATTCACATTTATGAGTTTTCAG 614
QY      194 ATCTGAAGACTCTAACCTTTGGCTTCATTCACCTTGCATTTACATGTGGCGCATCTTC 253
Db      615 CTCAGAGCAGATATATGTCATGCTGTTATAGCGGGTTGCATTTTTCATCTT 674
QY      254 TTGGCGGTGACGCACTTGTGAGGCTTGGCGCTGG 290
Db      675 TTTTACTTGCTTTTGATGCTGAAGCTAGGCTGGGG 711

```

RESULT 15

```

CK295424      936 bp      mRNA      linear      EST 15-DEC-2003
LOCUS      CK295424
DEFINITION      EST158138 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMCU77 5'
end, mRNA sequence.

```

```

ACCESSION      CK295424
VERSION      CK295424.1 GI:39879801
KEYWORDS
SOURCE      EST.
ORGANISM      Nicotiana benthamiana

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

```

```

REFERENCE      1 (bases 1 to 936)
AUTHORS      Buell, C.R., Hart, A., Zismann, V., Karamecheva, S.A., Day, B.,
Staskawicz, B., Jin, H. and Baker, B.

```

```

TITLE      Generation of EST sequences from Nicotiana benthamiana
JOURNAL      Unpublished (2003)
COMMENT      Other ESTs: EST738139

```

```

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-aray@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: RTT TNG GTG ACA CTA TAG.
Location/Qualifiers

```

FEATURES

```

source
1..936
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NBMCU77"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TOM"
/lab_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="Vector: PCW/SPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN

```

Query Match      10.7%; Score 34.6; DB 14; Length 936;
Best Local Similarity 47.5%; Pred. No. 26;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

```

```

QY      74 TTGCGTTCTTAATCCCTACGTAATCACATTTTGGCAAAATGGCATGT 133
Db      419 TTCTAATGTTAATGGTCAGCCGAAACATCTCAAGTGGCGGAAATTCGCGCTGT 478
QY      134 GGATGCCAGCAGAAATCTTGCCATTTTGAACCGCATCAAGTTTGCAGCATGATAG 193
Db      479 GGATGATACCAAGTGTTCATTTGCTTTGGAATTCCTCAATTATGAAGTTTTCAGAG 538

```

```

QY      194 ATCTGAAGACTCTAACCTTTGGCTTCATTCACCTTGCATTTACATGTGGCGCATCTTC 253
Db      539 CTCAGAGCAGATATATGTCATGCTGTTATAGCGGGTTGCATTTTTCATCTT 598
QY      254 TTGGCGGTGACGCACTTGTGAGGCTTGGCGCTGG 290
Db      599 TTTTACTTGCTTTTGATGCTGAAGCTAGGCTGGGG 635

```

```

Search completed: April 17, 2004, 20:01:47
Job time : 1053.18 secs

```


QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTGATACAAATTGGCGAAATGGCGATGT 133
 Db 404 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAGGTGGCGGAAATTCGCGCTGT 463
 QY 134 GGATGCCAGCAGGATCTCTTGGCATTGGACGCATCAAGCTTTCGGAGCAATGCCATAG 193
 Db 464 GGATGATACCAAGATGTTCATTTGCTTTGGAAATCCCAATTAAGAAATTTGGCAAG 523
 QY 194 ATCTGAAGACTTAACCTTTGATCTCATTTGCCGCTTGCATTACAGTGTGGCGCATCTTC 253
 Db 524 CTCAGAGCAAGATATAGTGTATAGCGTGTATAGCGGGGTTCGATTTGTTGCACTACT 583
 QY 254 TTGGCGGTGACGACCACTTTGTTGAGCGTTGGCGCTGG 290
 Db 584 TTTTACTTGGCTTTTGATGCTGAAGCTAGGCTGGGG 620

RESULT 13
 CK288648
 LOCUS 890 bp mRNA linear EST 15-DEC-2003
 DEFINITION Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NBMX102 5' end, mRNA sequence.

ACCESSION CK288648
 VERSION CK288648.1 GI:39866370
 KEYWORDS
 SOURCE
 ORGANISM
 Nicotiana benthamiana
 Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 890)
 Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A., Day, B., Staskawicz, B., Jin, H., and Baker, B.,
 Generation of EST sequences from Nicotiana benthamiana Unpublished (2003)
 CONTACT: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 location/Qualifiers
 1..890
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMX102"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_host="DH10B-TonA"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 890;
 Best Local Similarity 47.5%; Pred. No. 25;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTGATACAAATTGGCGAAATGGCGATGT 133
 Db 543 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAGGTGGCGGAAATTCGCGCTGT 602

QY 134 GGATGCCAGCAGGATCTCTTGGCATTGGACGCATCAAGCTTTCGGAGCAATGCCATAG 193
 Db 603 GGATGATACCAAGATGTTCATTTGCTTTGGAATTCCTCAATTAAGAAATTTGGCAAG 662
 QY 194 ATCTGAAGACTTAACCTTTGATCTCATTTGCCGCTTGCATTACAGTGTGGCGCATCTTC 253
 Db 663 CTCAGAGCAAGATATAGTGTATAGCGTGTATAGCGGGGTTCGATTTTGGCACTT 722
 QY 254 TTGGCGGTGACGACCACTTTGTTGAGCGTTGGCGCTGG 290
 Db 723 TTTTACTTGGCTTTTGATGCTGAAGCTAGGCTGGGG 759

RESULT 14
 CK295854
 LOCUS 922 bp mRNA linear EST 15-DEC-2003
 DEFINITION Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NBMX72 5' end, mRNA sequence.

ACCESSION CK295854
 VERSION CK295854.1 GI:3980657
 KEYWORDS
 SOURCE
 ORGANISM
 Nicotiana benthamiana
 Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 922)
 Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A., Day, B., Staskawicz, B., Jin, H., and Baker, B.,
 Generation of EST sequences from Nicotiana benthamiana Unpublished (2003)
 CONTACT: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source
 location/Qualifiers
 1..922
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMX72"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_host="DH10B-TonA"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 922;
 Best Local Similarity 47.5%; Pred. No. 26;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTGATACAAATTGGCGAAATGGCGATGT 133
 Db 495 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAGGTGGCGGAAATTCGCGCTGT 554
 QY 134 GGATGCCAGCAGGATCTCTTGGCATTGGACGCATCAAGCTTTCGGAGCAATGCCATAG 193

ORIGIN

/clone="CSODI052YB18"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 10.7%; Score 34.8; DB 13; Length 1201;
 Best Local Similarity 39.9%; Pred. No. 25; Indels 0; Gaps 0;
 Matches 63; Conservative 25; Mismatches 70; Gaps 0;

154 GCCATTTGACCGCATCAACGTTTGCAGCAATGGAGATAGATCTTAACCTT 213
 943 GACTTTCWCGAGCGCGCGCGTGTGTCGCGGAGCGCGCCCTTCCCGAGT 1002
 214 GGTCTCATTTGCGGTTGCGATTCAGTGTGGCGCATCTTGTGGAGCGACCTTG 273
 1003 GATKGGGTTGCGCGCTGGACCACTGTTTGTGGACT 311
 274 TTGAGCGTTGCGCGCTGGACCACTGTTTGTGGACT 311
 1063 CAGKTAGTGGGGGGGTSACVATCTTCGTSTTAAAGT 1100

RESULT 11 413 bp mRNA linear EST 30-APR-2003
 LOCUS CB239541
 DEFINITION BSH13B1 two-month-old roots from clone 'Beaurepe' grown for 19 days
 under restricted irrigation Populus balsamifera subsp. trichocarpa
 x Populus deltoides cDNA 5', mRNA sequence.

ACCESSION CB239541
 VERSION CB239541.1 GI:30233030
 KEYWORDS EST.
 ORGANISM Populus balsamifera subsp. trichocarpa x Populus deltoides
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosid1; Malpighiales; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 413)
 AUTHORS Kohler, A., Delaruelle, C., Martin, D. and Martin, F.
 TITLE The poplar root transcriptome: analysis of 7000 expressed sequence
 tags

JOURNAL Unpublished (2003)
 COMMENT Contact: Martin FM
 Equipe de Microbiologie Forestiere
 Institut National de la Recherche Agronomique
 Centre INRA de Nancy, 54280 Champenoux, France
 Tel: +33 383 39 40 80
 Fax: +33 383 39 40 69
 Email: fmarc@nancy.inra.fr

FEATURES
 source Seq primer: Fornat 5' AAGCGGCGCATTTGTGTGACCC.
 Location/Qualifiers

1. 413
 /organism="Populus balsamifera subsp. trichocarpa x
 Populus deltoides"
 /mol_type="mRNA"
 /cultivar="Beaurepe"
 /db_xref="taxon:3695"
 /dev_stage="two-month-old"
 /clone_id="two-month-old roots from clone 'Beaurepe' grown
 for 19 days under restricted irrigation"
 /note="Organ: root; Vector: pTRIPLEX2; cDNA library of
 roots from two-month-old Populus trichocarpa Torr. & Gray
 x deltoides Barts. Ex Marshall (clone 'Beaurepe') grown for
 19 days under restricted irrigation to reach 50% of the
 transpiration rate of fully watered plants. The cDNA
 library was constructed from 1 ug of total RNA using the
 SMART cDNA synthesis kit (Clontech, Palo Alto, CA, USA)
 according to the manufacturer's instructions. The
 resulting cDNA was packed into lambda phages using the

ORIGIN

Gigapack III Gold packaging kit (Stratagene, La Jolla,
 CA). The pTRIPLEX2 phagemid clones in Escherichia coli
 were obtained by using the mass in vivo excision protocol
 according to the manufacturer's instructions (Clontech)."

Query Match 10.7%; Score 34.6; DB 14; Length 413;
 Best Local Similarity 60.4%; Pred. No. 19;
 Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

65 TCCGGGCGGTTCCGTTCTTAATCCTTACGCCCTACGTAATACATTTGGGCAAAA 124
 71 TCCGAGGCTCCCTCTGAGAGCTTTTACCCTCCAGAGATTCACCTTCAAGTCAAA 130
 125 TGGCGATGTGATGCGAGCAGAACTTGC 155
 131 TGGGGTTCCGGCGGCGCACCGAGTTCCTGAC 161

RESULT 12 853 bp mRNA linear EST 15-DEC-2003
 LOCUS CK297477
 DEFINITION EST760191 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NBMD912 5'
 end, mRNA sequence.

ACCESSION CK297477.1 GI:39883902
 VERSION CK297477
 KEYWORDS EST.
 ORGANISM Nicotiana benthamiana
 Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1 (bases 1 to 853)
 AUTHORS Buell, C.R., Hart, A., Zisman, V., Karanicheva, S.A., Day, B.,
 Stekawicz, B., Jin, H. and Baker, B.
 TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST760192
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source

1. 853
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMD912"
 /tissue_type="biotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_id="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1, Site 1: EcoRI, Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 853;
 Best Local Similarity 47.5%; Pred. No. 25; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, Z., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, U., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

1. 728
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F830034E14"
/issue_type="activated spleen"
/clone_lib="RIKEN full-length enriched, activated spleen"

ORIGIN

Query Match 10.7%; Score 34.8; DB 13; Length 728;
Best Local Similarity 55.9%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 188 CGATGATCTGAAGACTTACCTTGTGCTCTATTCGCTGCGATTAAGTGTGGCGC 247
DB 244 CGACGAGCTGATCTTTATCCCTTGTGTAGCTGCTGTTAGACAGAGCTGGTGC 303
QY 248 ATTTCTTGGCGGTGACGACCTTGTGTAGCGTTGGCGCTGGACCACTGTTTGT 305
DB 304 AACATCTGTGTTCGACGCCCTCTCTTGTGCGTGTGTCTAAGGCACTGCTTGT 361

RESULT 9
LOCUS B1144284 808 bp mRNA linear EST 05-JUL-2001
DEFINITION 60290725F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5064373
5', mRNA sequence.
ACCESSION B1144284
VERSION B1144284.1 GI:14604285
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 808)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1173 row: m column: 14
High quality sequence start: 6
High quality sequence stop: 793.

FEATURES

source

1. 808
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064373"
/db_xref="IMAGE:5064373"
/clone_lib="NCI_CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 10.7%; Score 34.8; DB 12; Length 808;
Best Local Similarity 53.7%; Pred. No. 21;
Matches 72; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 133 TGAATGACAGAGATACCTTGCATTTTGACCGCATCAAGCTTTGCGACATGAGATA 192
DB 523 TGGAGGCTTCATGCTTCAACTTATGTGACGACCACTGGACATCAACAGACATG 562
QY 193 GATCTGAAGACTCTAACCTTGTGCTCTATTCGCTGCGATTAAGTGTGGCGCATCTT 252
DB 563 GTTCAGAGATAGAGAGCTTGGGTTTCAAGCTTGTGTGCTCATGTAGATGCGCTGTA 642
QY 253 CTGGCGGTGACG 266
DB 643 CTGGCAATAGGCG 656

RESULT 10
LOCUS BX387771 1201 bp mRNA linear EST 08-MAY-2003
DEFINITION BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1052YB18 5-PRIME, mRNA sequence.
ACCESSION BX387771
VERSION BX387771.1 GI:30460721
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4682.r For more information about this cluster, see <http://www.genoscope.cns.fr/>
cgi-bin/cluster.cgi?seq=CS0A1052DA09QPl&cluster=4682.r. Contact : Feng Huang Email : Fluang@litech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Paraday Avenue Genoscope sequence ID : CS0A1052DA09QPl.

FEATURES

source

1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN

/strain="B73"
/db_xref="taxon:4577"
/clone="ZM07A530007"
/note="Vector: pCR4-TOPO, Site 1: EcoRI, 0.6-1.0 kb high
cot selected genomic DNA library"

Query Match 10.9%; Score 35.2; DB 28; Length 910;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 192 AGATCTGAAGACTTAACTTGTGCTCATTCCTGCGATTAACAGTGGCGGCACT 251
DB 272 AATCTGACGGCTCAATCTCTTCCCAATCCGACATCGGACAGGCGCGCGCC 331
QY 252 TCTTGCGCGCTGACGACCTTGTGAGCGTTGCGCGTGGACCATGTTTT 303
DB 332 TGGTGGCGCGAATCTCTGCTGCGACGACCGCATGCTGCGCGCGTTT 383

RESULT 7 1101 bp DNA linear GSS 04-JUN-1999
CNS00BD8
LOCUS Drosophila melanogaster genome survey sequence TETS end of BAC #
DEFINITION BAC23H11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL056828 GI:4937496
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
The BAC end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using the BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Oosawa and
Aaron Mammeter in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source
1. 1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR23H11"
/clone_11b="RPCI-98"
/note="end : TETS"

ORIGIN

Query Match 10.9%; Score 35.2; DB 29; Length 1101;
Best Local Similarity 24.4%; Pred. No. 18;
Matches 62; Conservative 81; Mismatches 111; Indels 0; Gaps 0;

QY 29 TTGTTGCGCAGATGTCAGATCATTTCTTGGCGTCCGCGGCGCTTCCTTAATCC 88
DB 847 TTGGGKTGGTGGGTAASDTGTTATATTTTBTBKBSTBTBKAATKWSASSTTT 906

QY 89 TTAGCCCTTACGTATCATCTTTTGCGCAAAATGCGCATGTGCAGCAGGAA 148
DB 907 WSGAGGSGTSTKTBKSGTSTSVGTTTWSGSAVAASGTTTWTSTBTKAABBA 966
QY 149 TCCTGCATTTTGACCGCATCAACGTTTGCAGCATGCGATAGATCTGAAGACTTAA 208
DB 967 ATGASHTTASDPTWAABSSKTSVBSBSSDMGCTGASGATTTTCTTTTWTWW 1026
QY 209 CTTTGTCTCATTCGCGTGGCATTAACGTGCGCATCTTCTGCGCGTCCAGCA 268
DB 1027 TATTNTSBBBWBATTTTDTAAATWAATBSTNSNNMTTBTSSRAATRAABVA 1086

RESULT 8 728 bp mRNA linear EST 17-DEC-2002
BY765711
LOCUS cDNA clone F830034E14 3', mRNA sequence.
DEFINITION
ACCESSION BY765711 GI:27203910
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE

AUTHORS
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Otsu, N., Saito, R., Suzuki, H., Yamataka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Balt, C.,
Hume, D.A., Quakenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V.,
Choitha, C., Corban, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Matsuda, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, I., Pavan, W.J., Pette, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilmberg, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, T.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Cantinici, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Komori, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arikawa, T., Fukuda, S., Hara, A., Hashizume, M., Imutani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22346851
12466851
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9216
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P.,

LOCUS CF38060 475 bp mRNA linear EST 18-AUG-2003
 DEFINITION JMT--08-M05.b1 AtUMT-overexpressing transgenic rice plasmid cDNA
 library (JMT) Oryza sativa cDNA clone JMT--08-M05, mRNA sequence.
 ACCESSION CF38060
 VERSION CF38060.1 GI:33824491
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 475)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,Y.K., Kim,Y.-K. and Nahm,B.H.,
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 TITLE JOURNAL
 COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
 Location/Qualifiers
 1..475
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:4530"
 /clone="JMT--08-M05"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="AtUMT-overexpressing transgenic rice plasmid
 cDNA library (JMT)"
 /note="Vector: PCR4-TOPO, Site 1: EcoRI; Oligo-capped mRNA
 was reverse transcribed and then used for PCR. mRNA was
 prepared from Arabidopsis thaliana Uasmonte Carboxyl
 methyltransferase overexpression line."

ORIGIN
 Query Match 10.9%; Score 35.2; DB 14; Length 475;
 Best Local Similarity 50.6%; Pred. No. 13; Indels 0; Gaps 0;
 Matches 85; Conservative 0; Mismatches 83;
 51 CATTACTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTGAATCA 110
 DB 255 CTTTCTTTCTGGGCTCCCGGGGCTTCTGCTAGCCTTCTTAACGCGCCCTTGGCGGCG 196
 QY 111 ATTTCGGGCAAAATGGGCAATCTGATCCAGCGAGATCCTTGCACTTTGACCGCATC 170
 DB 195 CGGCGCTTTCGGCGCTCTACTAGCTTGCGACCGCGCTCTTGGAAGTCCGCGAGGTTT 136
 QY 171 AACGTTTGGCAGCATGCGATGATCTGAAGCTCTTAACCTTTGGTCT 218
 DB 135 CACGCGCTTCAGAGAGCATCTTCTGATGCTTCTGCTGCTGTTCT 88

RESULT 5
 B2720836 786 bp DNA linear GSS 24-FEB-2003
 LOCUS PUCFMO5TD ZM 0.6.1.0 KB Zea mays genomic clone ZMABTA137A09,
 DEFINITION genomic survey sequence.
 ACCESSION B2720836
 VERSION B2720836.1 GI:28512409
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogonaeae; Zea.
 REFERENCE 1 (bases 1 to 786)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,

RENNICK,A., FRASER,C.M., YUAN,Y., SAN MIGUEL,P., MA,J. and
 BENNETZEN,J.
 Maize Genomics Consortium
 Unpublished (2003)
 COMMENT Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..786
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMABTA137A09"
 /clone_lib="ZM 0.6.1.0 KB"
 /note="Vector: PCR4-TOPO, Site 1: EcoRI; 0.6-1.0 kb high
 cot selected genomic DNA library"

ORIGIN
 Query Match 10.9%; Score 35.2; DB 28; Length 786;
 Best Local Similarity 53.7%; Pred. No. 16; Indels 0; Gaps 0;
 Matches 73; Conservative 0; Mismatches 63;
 189 GATAGATCGAAGACTCTAACCTTGTCTCATTGCGCGTGGAGTATACAGTGGTGGCGCA 248
 DB 480 GAAGCAATCAAAATAATTGTGTCTAGTCTCAAGGTGGTGGCAATCTGTGTGGCTTA 539
 QY 249 TCTTCTTGGCGGTGACGACCTTGTGAGCGTGTGGCGCTGACCAATCTTTTGTGG 308
 DB 540 TCATTTAGGTGTCATCTTAGAATGTTATCATTTGGCCCGCCAGCAATCTTGTGATCG 599
 QY 309 ACTGTGATCTTTTC 324
 DB 600 TATGTTGCACATGTGC 615

RESULT 6
 CC362201 910 bp DNA linear GSS 16-MAY-2003
 LOCUS PHTK88TD ZM 0.6.1.0 KB Zea mays genomic clone ZMABTA1330007,
 DEFINITION genomic survey sequence.
 ACCESSION CC362201
 VERSION CC362201.1 GI:30831601
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogonaeae; Zea.
 REFERENCE 1 (bases 1 to 910)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Renwick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 COMMENT Other GSSs: PHTK88TB
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..910
 /organism="Zea mays"
 /mol_type="genomic DNA"

FEATURES
 source

QY 178 CCAGCAATGATAGATCTGAAGACTTAACCTTTGCTCATTCGCGTTCGATTTACA 237
DB 15721 GGCTCAGCTGCGGGGAAGAGCCCTTGAAGCAGACGACCATCTGCTGTGCGGTGTGA 15780
QY 238 GTGGTGGCGCATCTTTCTTGGCG 259
DB 15781 GTGGTACTCTATCTTCTGTGTG 15802

RESULT 14

US-10-796-280-12378
; Sequence 12378, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12378
; LENGTH: 234309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(234309)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-796-280-12378

Query Match 8.8%; Score 28.4; DB 6; Length 234309;
Best Local Similarity 52.5%; Pred. No. 81;
Matches 62; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 206 TAACCTTTGCTCATTCGCGTTCGATTAACAGTGGCGCATCTTTCGCGGTGAC 265
DB 62038 TCACCTCAGGCGCTTTAGCCTGAAGAGCGTGGCCCTACATCTGCTTTAGTTTCA 62097
QY 266 GCACTTTGTTGAGCGTGGCGCTGGCAGCATGTTTGTGGACTGGTAATCTTTT 323
DB 62098 GTCCCTGTGGTAAGTTGGCTAAGGCCCTTCCCTGTGTGTAAGAGGGGTTT 62155

RESULT 15

US-10-767-701-27384/C
; Sequence 27384, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535) B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 27384
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 6858297
US-10-767-701-27384

Query Match 8.7%; Score 28.2; DB 6; Length 490;
Best Local Similarity 51.2%; Pred. No. 5.8;
Matches 66; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 192 AGATCTGAAGACTTAACCTTTGCTCATTCGCGTTCGATTAACAGTGGCGCATCT 251
DB 282 AGGTGTACAGTGTAGCAGTTGTGTACTGCGCAGAGCAGCGAGTGAAGCCAGCGCC 223

QY 252 TCTTGGCGGTGACAGCAGCACTTGTGAGCGTTGGCGCTGGCACCAGCATGTTTTGTGACT 311
DB 222 AGTTGGCGGTGAGAGGTGACCGACTGGGTGGCGGTGTGCAACCAAGACTTGTGTCACT 163
QY 312 GTGATCT 320
DB 162 GGGCAACT 154

Search completed: April 17, 2004, 20:13:00
Job time : 39.4407 secs

QY 143 CAGGATCTTCCATTTTGACCCGATCAGCTTTCGAGCAAGTCCGATAGATCTGAGA 202
DB 284 CCGGATCTTCCGCGGAGCTATACCGACGCTATTAACCGGATGCTTCCCTGCCACGG 225
QY 203 CTCTAACCCTTGGTCTCATTTGCGTTGCGATTACAGTGGTGGCGCATCTTTCGCGGTC 262
DB 224 CACTTAACCGTACGGGTATTTGGCTTGGCCGACCACTGATTGCGCTTCAAGTGGCGATG 165
QY 263 GACGACCTTGTG 276
DB 164 TTCTGATCATGATG 151

RESULT 10
US-10-108-605A-248/C
; Sequence 248, Application US/10108605A
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Steam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kander, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 3113B
; CURRENT APPLICATION NUMBER: US/10/108,605A
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 248
; LENGTH: 8577
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605A-248

Query Match 8.8%; Score 28.4; DB 6; Length 8577;
Best Local Similarity 60.3%; Pred. No. 18;
Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 209 CCTTGATCTTCATTCGCGCTTGCGATTAAGTGGTGGCGCATCTTTCGCGGTCACGCA 268
DB 83 CCGTGTGTTGTTGGCGCCACTGCTACTGTGTCTGCTTGGCAGGCGACGAC 24
QY 269 CCTTGTGAGCGCTTGGCG 286
DB 23 CCCGTTCAACGGCGGTG 6

RESULT 11
US-10-796-280-12417/C
; Sequence 12417, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12417
; LENGTH: 20618
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12417

Query Match 8.8%; Score 28.4; DB 6; Length 20618;
Best Local Similarity 51.6%; Pred. No. 27;
Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGGGGGTTCGGTCTTAATCCTTAAGCCCTAGCATGATCAGAAATTTGGGCAAAA 124
DB 8143 TCCCATCATTTCCACTGTTATACCTTTGGCTCAAGCAGACATCACCTGTCTTGATTTA 8084
QY 125 TGGCGATGTGATGACGAGCAAGAAATCTTGCAATTTTGAACCGCATCAACGTTTCGAGCA 184
DB 8083 TGCTGATCTACTGCTTCATCTCTGCTGCGCTGTCTTCCCACTATTTCTACCA 8024
QY 185 ATGGCA 190
DB 8023 AGGCA 8018

RESULT 12
US-10-796-280-12418
; Sequence 12418, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12418
; LENGTH: 46075
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12418

Query Match 8.8%; Score 28.4; DB 6; Length 46075;
Best Local Similarity 51.6%; Pred. No. 39;
Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGGGGGTTCGGTCTTAATCCTTAAGCCCTAGCATGATCAGAAATTTGGGCAAAA 124
DB 3888 TCCCATCATTTCCACTGTTATACCTTTGGCTCAAGCAGACATCACCTGTCTTGATTTA 3947
QY 125 TGGCGATGTGATGACGAGCAAGAAATCTTGCAATTTTGAACCGCATCAACGTTTCGAGCA 184
DB 3948 TGCTGATCTACTGCTTCATCTCTGCGCTGCGCTGCTTCCCACTATTTCTACCA 4007
QY 185 ATGGCA 190
DB 4008 AGGCA 4013

RESULT 13
US-10-796-280-12452
; Sequence 12452, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12452
; LENGTH: 69359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-12452

Query Match 8.8%; Score 28.4; DB 6; Length 69359;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 118 GGCAGAAATGCGGATGTGATGCCAGAGAAATCTTGCAATTTTGAACGATCAACGTTT 177
DB 15661 GGCAGAAATGCGGATGTGAAATAGCAGATTGACTCCAGAGATCTGTGCCCCCTGTCCACTT 15720

QY 108 ACAATTGTGGCAAAATGCGGATGTGATGCCAGCAGA 147
DB 48937 ACAAGTTGAATCCAAATGCCCAAGGTATGCGCTTGGGA 48898

RESULT 6
US-10-650-609-1/c

Sequence 1, Application US/10650609
GENERAL INFORMATION:
APPLICANT: Costa, Robert
APPLICANT: Wang, Xinde
APPLICANT: Tan, Yongjun
APPLICANT: Kalinichenko, Vladimir
APPLICANT: Krupczak-Holts, Katherine
APPLICANT: Wang, I-Ching
APPLICANT: Major, Michael
TITLE OF INVENTION: METHODS OF TREATING AGE-RELATED DEFECTS AND DISEASES
FILE REFERENCE: 02-1040-A
CURRENT APPLICATION NUMBER: US/10/650,609
CURRENT FILING DATE: 2003-08-28
PRIOR APPLICATION NUMBER: 60/406582
PRIOR FILING DATE: 2003-08-28
PRIOR APPLICATION NUMBER: 60/426068
PRIOR FILING DATE: 2002-11-13
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 2737
TYPE: DNA
ORGANISM: Homo sapiens
US-10-650-609-1

Query Match 8.8%; Score 28.6; DB 6; Length 2737;
Best Local Similarity 55.6%; Pred. No. 9.2;

Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 61 GCGCTCGGCGGCGGCTTCCCTTAATCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
DB 1109 GAGCTCTGATTCGATCGCTTCTGCTGCTTAACACCTGTCATGCAAGTACGCGGTT 1050
QY 121 AAATGCGCATGTGATGCGACAGCAATCCTTGCCATT 159
DB 1049 GGCACCTGGGATGATGCTCCAGAGAGACCTTGCCATT 1011

RESULT 7
US-10-451-467A-545/c

Sequence 545, Application US/10451467A
GENERAL INFORMATION:
APPLICANT: CONTRERAS, ROLAND HENRI
APPLICANT: EBERHARDT, INES
APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
APPLICANT: REEKMAN, RIEKA JOSEPHINA
TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
FILE REFERENCE: JAB-1667
CURRENT APPLICATION NUMBER: US/10/451,467A
CURRENT FILING DATE: 2003-06-19
PRIOR APPLICATION NUMBER: EP 00870318.3
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: EP 01870002.1
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: EP 01870003.9
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 732
SOFTWARE: PatentIn version 3.1
SEQ ID NO 545
LENGTH: 3042
TYPE: DNA
ORGANISM: Candida albicans
US-10-451-467A-545

Query Match 8.8%; Score 28.6; DB 6; Length 3042;
Best Local Similarity 59.0%; Pred. No. 9.6;
Matches 49; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 9 TGAATTCCTCGTATTCCTTGTGTGCGAGTATGCAATCTTACTTTGCGCTCCG 68
DB 956 TGTGCTCTTTCATCTCTTGTGTTCAATTTGTTCTTTCTTTGATGATTT 897
QY 69 GCGGCTTCGCTTATCCTTA 91
DB 896 TGTGTTCCAAATTTGATTTTAA 874

RESULT 8
US-10-767-701-21673/c

Sequence 21673, Application US/10767701
GENERAL INFORMATION:
APPLICANT: KOVALIC, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 21673
LENGTH: 312
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 13317221
US-10-767-701-21673

Query Match 8.8%; Score 28.4; DB 6; Length 312;
Best Local Similarity 53.6%; Pred. No. 4;
Matches 59; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTCATTCGCTTGCGATTCACAGTGTGGCGCATCTTGTGGCGGTG 263
DB 261 TCTTATCTCTTCCTGCGCTTGCCTTCCTGCTGCTGCGCTTCTTCCGCGGAGG 202
QY 264 ACGACCTTGTGAGCGTGGCGCTGCGACCAATCGTTTGTGACTGG 313
DB 201 GCGCGAGGAGCGCGCGCTGCGCGCTGCGCGCGGCGGCGGATGGAGAGCGGG 152

RESULT 9
PCT-US04-02000-767/c

Sequence 767, Application PC/TUS0402000
GENERAL INFORMATION:
APPLICANT: Science Applications International Corporation
TITLE OF INVENTION: Method and System for Identifying Biological Entities in
FILE REFERENCE: 36609-183264 (SAIC0087-PCT)
CURRENT APPLICATION NUMBER: PCT/US04/02000
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US 60/441,745
PRIOR FILING DATE: 2003-01-23
PRIOR APPLICATION NUMBER: US 60/441,806
PRIOR FILING DATE: 2003-01-23
NUMBER OF SEQ ID NOS: 3242
SOFTWARE: PatentIn version 3.2
SEQ ID NO 767
LENGTH: 500
TYPE: DNA
ORGANISM: Escherichia coli
PCT-US04-02000-767

Query Match 8.8%; Score 28.4; DB 1; Length 500;
Best Local Similarity 50.7%; Pred. No. 4.9;
Matches 68; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

US-10-789-400-36
Sequence 36, Application US/10789400
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
APPLICANT: HUMAN SERVICES
APPLICANT: COLLINS, Peter L.
APPLICANT: Biacchesi, Stephanie
APPLICANT: Buchholz, Ursula
APPLICANT: Skladopoulos, Mario H.
APPLICANT: Murphy, Brian R.
TITLE OF INVENTION: RECOMBINANT HUMAN METAPNEUMOVIRUS AND ITS USE
FILE REFERENCE: 4239-67783
CURRENT APPLICATION NUMBER: US/10/789,400
CURRENT FILING DATE: 2004-02-27
PRIOR APPLICATION NUMBER: US 60/451,119
PRIOR FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/478,667
PRIOR FILING DATE: 2003-06-13
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.2
SEQ ID NO 36
LENGTH: 13350
TYPE: DNA
ORGANISM: human metapneumovirus
US-10-789-400-36

Query Match 9.4%; Score 30.4; DB 6; Length 13350;
Best Local Similarity 49.4%; Pred. No. 4.5; Mismatches 81; Indels 0; Gaps 0;
Matches 79; Conservative 0;

QY 14 TCCTCGATATCTCTCTGTTGCGAGATGTCAGTCACTTATCTTTGGCTCCGGGCGG 73
DB 1687 TCTCATATATGAGAGAGAGATGCGATCTTCAATCTTAACTTTGAGAGAGATA 1746
QY 74 TTCCTGTTTATCCCTTAGCCCTCGATGATCAATTTTGGGCAAAATGGGATGT 133
DB 1747 CTTGATCATTTAGCATTTAGGCGGAGATTTGATTAATTAAGAGAAATTTAGCATGAT 1806
QY 134 GGATGCCAGCAGGATCTTGCATTTTACCGCATCAAC 173
DB 1807 TAGGGCTATTTAGACATCACTGATTCAGCAGGACC 1846

RESULT 3
US-10-767-471-10749/C
Sequence 10749, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10749
LENGTH: 35895
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(35895)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10749

Query Match 9.3%; Score 30.2; DB 6; Length 35895;
Best Local Similarity 52.8%; Pred. No. 8.2; Mismatches 58; Indels 0; Gaps 0;
Matches 65; Conservative 0;

QY 201 GACTCTAACCTTGTCTCATTTGCGGATTTAGCATTAAGTGTGCGGCGCATCTTTGGCGG 260
DB 12662 GCTCTACACCATGCTCTCTTGGGTTTGTCTTGGCAGCAAGAGCCATCTCTGCGCTG 12603

QY 261 TCGACGACCTTGTGAGCGTTGCGGCTGCGACCATCTTTTGTGACTGGTGAATCT 320
DB 12602 CCGACATCTGCTTGGCTCTCGAGCTCCGACGATGTTTGACACTGACACT 12543
QY 321 TTT 323
DB 12542 TCT 12540

RESULT 4
US-10-767-701-5423/C
Sequence 5423, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5423
LENGTH: 731
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS132149_1
US-10-767-701-5423

Query Match 8.9%; Score 28.8; DB 6; Length 731;
Best Local Similarity 53.6%; Pred. No. 4.3; Mismatches 52; Indels 0; Gaps 0;
Matches 60; Conservative 0;

QY 149 TCCTTGCATTTTGAACCGCATCAAGTTTGCAGCAATGGATGATCTGAAGACTCTAA 208
DB 240 TTTCGGCAGATTGTCTCTCGGCGACACATCCACATTTGATTGTTTCAAGACTAAATG 181
QY 209 CTTTGTTCATTTGCGGCTTGCATTTACATGTTGGCGCATCTTTGGCGG 260
DB 180 CATCTCATCATCTTGAAGTTGCGAGTTAGAGCTGAACCACTCTCGGCGTG 129

RESULT 5
US-10-767-471-10669/C
Sequence 10669, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10669
LENGTH: 158027
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(158027)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10669

Query Match 8.9%; Score 28.8; DB 6; Length 158027;
Best Local Similarity 54.0%; Pred. No. 50; Mismatches 44; Indels 0; Gaps 0;
Matches 54; Conservative 2;

QY 48 AGTCATTAATCTTTGCGGCTCCGGGCGGCTTCGTTCTTAATCTTTAAGCCCTTACGTGATC 107
DB 48997 AGTCATTAATCTTTAAGCCAGAGCTGTATGATGATGATCTGTGTCCTTCAAAATTC 48938

Tue Apr 20 06:47:20 2004

us-10-608-504-4.rmpn

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 37.4407 Seconds
(without alignments)
5694.647 Million cell updates/sec

Title: US-10-608-504-4
Perfect score: 324
Sequence: 1 atgacactgattctctctcctcgttgcgtggtgacatcttttc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1362413 seqs, 329030602 residues

Total number of hits satisfying chosen parameters: 2724826

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Pending Patents NA New.*
1: /cgn2_6/prodata/1/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/prodata/1/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/prodata/1/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/prodata/1/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/prodata/1/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.2	10.2	173233	6 US-10-021-698A-716	Sequence 716, App
2	30.4	9.4	13350	6 US-10-789-400-36	Sequence 36, Appl
3	30.2	9.3	35895	6 US-10-767-471-10749	Sequence 10749, A
4	28.8	8.9	731	6 US-10-767-701-5423	Sequence 5423, Ap
5	28.8	8.9	158027	6 US-10-767-471-10669	Sequence 10669, A
6	28.6	8.8	2737	6 US-10-650-609-1	Sequence 1, Appl
7	28.6	8.8	3042	6 US-10-451-467A-545	Sequence 545, App
8	28.4	8.8	312	6 US-10-767-701-21673	Sequence 21673, A
9	28.4	8.8	500	6 PCT-US04-02000-767	Sequence 767, App
10	28.4	8.8	8577	6 US-10-108-605A-248	Sequence 248, App
11	28.4	8.8	20618	6 US-10-796-280-12417	Sequence 12417, A
12	28.4	8.8	46075	6 US-10-796-280-12418	Sequence 12418, A
13	28.4	8.8	69359	6 US-10-796-280-12452	Sequence 12452, A
14	28.4	8.8	234309	6 US-10-796-280-12378	Sequence 12378, A
15	28.2	8.7	490	6 US-10-767-701-27384	Sequence 27384, A
16	28.2	8.7	531	6 US-10-767-701-25301	Sequence 25301, A
17	28.2	8.7	77294	6 PCT-US03-36500-1	Sequence 1, Appl
18	28.2	8.7	185248	6 US-10-796-280-12228	Sequence 12228, A
19	28	8.6	528	6 US-10-767-701-2761	Sequence 2761, Ap
20	28	8.6	593	6 US-10-767-701-29783	Sequence 29783, A
21	28	8.6	1102	6 US-10-021-698A-4460	Sequence 4460, Ap
22	28	8.6	1758	6 US-10-100-683-1035	Sequence 1035, Ap
23	28	8.6	2178	6 PCT-US04-05654-2655	Sequence 2655, Ap
24	28	8.6	4990	6 US-10-100-683-4553	Sequence 4553, Ap
25	27.8	8.6	1226	6 PCT-US04-09299-18	Sequence 18, Appl
26	27.8	8.6	1675	6 US-10-152-372-135	Sequence 135, App

27	27.8	8.6	1709	6 US-10-100-683-3274	Sequence 3274, Ap
28	27.6	8.5	201	6 US-10-767-471-27293	Sequence 27293, A
29	27.6	8.5	201	6 US-10-767-471-27441	Sequence 27441, A
30	27.6	8.5	201	6 US-10-796-307-18171	Sequence 18171, A
31	27.6	8.5	410	6 US-10-767-701-29976	Sequence 29976, A
32	27.6	8.5	35884	6 US-10-796-280-12303	Sequence 12303, A
33	27.6	8.5	37687	6 US-10-796-280-12200	Sequence 12200, A
34	27.6	8.5	417046	6 US-10-796-307-8720	Sequence 8720, Ap
35	27.4	8.5	509	6 US-10-767-701-18755	Sequence 18755, A
36	27.4	8.5	699	6 US-10-767-701-12271	Sequence 12271, A
37	27.4	8.5	1197	6 US-10-771-241-144	Sequence 144, App
38	27.4	8.5	1657	6 US-10-100-683-5402	Sequence 5402, Ap
39	27.4	8.5	76829	6 US-10-765-790-27	Sequence 27, Appl
40	27.2	8.4	600	7 US-60-545-213-175	Sequence 175, App
41	27.2	8.4	600	7 US-60-545-213-176	Sequence 4447, Ap
42	27.2	8.4	600	7 US-60-545-213-4447	Sequence 4447, Ap
43	27.2	8.4	600	7 US-60-545-213-4448	Sequence 4448, Ap
44	27.2	8.4	1365	6 US-10-417-884A-1659	Sequence 1659, Ap
45	27.2	8.4	1500	1 PCT-US04-09323-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1					
US-10-021-698A-716					
Sequence 716, Application US/10021698A					
GENERAL INFORMATION:					
APPLICANT: KEITH, TIM					
APPLICANT: LITTLE, RANDALL					
APPLICANT: VAN BERGEDEGH, PAUL					
APPLICANT: DUPUIS, JOSEF					
APPLICANT: DEL MASTRO, RICHARD					
APPLICANT: SIMON, JASON					
APPLICANT: ALLEN, KRISTINA					
APPLICANT: PANDIT, SONIL					
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO					
TITLE OF INVENTION: RESPIRATORY DISEASES AND OBESITY					
FILE REFERENCE: 2976-4044US1					
CURRENT APPLICATION NUMBER: US/10/021,698A					
CURRENT FILING DATE: 2001-10-22					
PRIOR APPLICATION NUMBER: 60/211,749					
PRIOR FILING DATE: 2000-06-14					
NUMBER OF SEQ ID NOS: 6160					
SOFTWARE: PatentIn 2.1					
SEQ ID NO 716					
LENGTH: 173233					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: modified_base					
LOCATION: (88987)..(89086)					
OTHER INFORMATION: a, t, c or g					
FEATURE:					
NAME/KEY: modified_base					
LOCATION: (167286)..(167385)					
OTHER INFORMATION: a, t, c or g					
US-10-021-698A-716					
Query Match					
Best Local Similarity 55.1%; Pred. No. 1.5;					
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;					
10.2%; Score 33.2; DB 6; Length 173233;					
QY 200 AGACTTACCTTGTCTATCCCTGCGATGAGTGGCGATCTCTGGCG 259					
DB 84488 AGAACTTACCTTGTCTATCCCTGCGATGAGTGGCGATCTCTGGCG 84547					
QY 260 GTGACGACCTTGTGAGCGTGGCGATGAGTGGCGATCTCTGGCG 317					
DB 84548 GTTACTTCAACCTCTGCTGCTGGCGATGAGTGGCGATCTCTGGCG 84605					
RESULT 2					

Tue Apr 20 06:47:19 2004

us-10-608-504-4.rni

Page 7

ADDRESS: Woodcock Washburn Kurtz Mackiewicz
ADDRESSER: No. 6031501st
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Loti Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439

INFORMATION FOR SEQ. ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-476-900A-4

	Query Match	9.2%	Score 29.8	DB 3	Length 5110
	Best Local Similarity	58.4%	Pred. No. 3		
	Matches	52	Conservative	0	Matches 37
					Indels 0
					Gaps 0
QY	118	GGCAAAATGGGCGATGTGGATGCGACAGAAATCCTTGCAATTTGACCGCATCAAGTTT	177		
Db	4500	GGCACTTGGCGCATGTCATGCGCATGATGATGATAGCGCTTCAAAATATCCGCATT	4441		
QY	178	CGCAGCAATGCGATGATCTGAAGACTCT	206		
Db	4440	CGGAAAAAGCGAGAGAGAGAGAGGATTT	4412		

Search completed: April 19, 2004, 18:38:15
Job time : 28.2807 secs

ADDRESS: Woodcock Washburn Kurtz Mackiewicz
ADDRESSER: No. 6031501st
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439

INFORMATION FOR SEQ. ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
HYPOTHEICAL: NO
ANTI-SENSE: NO

US-08-476-900A-4

	Query Match	9.2%	Score 29.8	DB 3	Length 5110
	Best Local Similarity	58.4%	Pred. No. 3		
	Matches	52	Conservative	0	Matches 37
					Indels 0
					Gaps 0
QY	118	GCGAATAATGGCGATGTGGATGCGACAGAAATCCTTGCAATTTGACCGCATCAAGTTT	177		
DB	4500	GCGACCTTGGCGCATGTCATGCGCATGATGATGATAGCGCCTTCAAAATATCCGCATT	4441		
QY	178	CGCAGCAATGCGATGATCTGAAGACTCT	206		
DB	4440	CGGAAAAAGCGAGAGAGAGAGAGGATTT	4412		

Search completed: April 19, 2004, 18:38:15
Job time : 28.2807 secs

US-08-488-546A-27

Query Match	9.2%;	Score 29.8;	DB 3;	Length 4635;
Best Local Similarity	58.4%;	Pred. No. 2.8;		
Matches 52;	Conservative	0;	Mismatches 37;	Indels 0;
				Gaps 0

QY	118	GGCAAAATGCGGATGTGGATGCGACACAGAAATCCTTGCCATTTTACCGACATCAACGTTT	177
Db	4025	GGCAGCTTGGCGCATCTCAATGCCATCGATGATGATACGCCCTTCAAAATATCCACCAAT	3966
QY	178	CGCAGCAATGGCATGTGATCTGAAGCTCT	206
Db	3965	CGGAAAAAGCCCGAGAGAGAGAGATTT	3937

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	9.2%;	Score 29.8;	DB 2;	Length 5110;
Best Local Similarity	58.4%;	Pred. No. 3;		
Matches	52;	Conservative	0;	Mismatches 37; Indels 0; Gaps 0;
Qy	118	GGCAAAATGGCCGATGCGATGCCAGCAGCAAACTCTTGCATTTGACCGATCAACGTTT	177	
Db	4500	GGCAGCTGGCGATGTCATGCAATGCCATGCATATGATAGCCCTTCAACATATCCACCAT	4441	
Qy	178	CGCAGCAATGCGATAGATCTGAAAGACTCT	206	
Db	4440	CGGAAAAAGCGCAGAGCAAGAGGATTT	4412	

```

RESULT 14
US-08-404-531B-5/C
; Sequence 5, Application US/08404531B
; Patent No. 5863724
; GENERAL INFORMATION:
; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
; APPLICANT: Thomas, Gilbert Cote, and Robert Gage!
; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonuryrea Receptor
; Patent No. 5863724
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSES: No. 5863724r1s
; STREET: One Liberty Place 46th. Floor
; CITY: Philadelphia
;

```

```

: COUNTRY: USA
: ZIP: 19103
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/404,531B
: FILING DATE: 15-MAR-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Beardsell, Lori Y.
: REGISTRATION NUMBER: 34,293
: REFERENCE/DOCKET NUMBER: BYLR-0003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 215-568-3100
: TELEFAX: 215-568-3439
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5110 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 260..5004
: US-08-404-531B-5

Query Match 9.2%; Score 29.8; DB 2; Length 5110;
Best Local Similarity 58.4%; Pred. No. 3;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 118 GGCAGAAATGGGGAGTGGATGCCAGCAGAGATCTTGGCATTTTGACCGCATCAAGCTT 177
DB 4500 GGCACCTTGGCGAGATCATATCCATCGATGATGATAGCGCCCTTCAACATATCCACATT 4441
QY 178 CGCAGCAATGCCATAGATCTTGAAGACTCT 206
DB 4440 CGGAAAAAGCGCAGAGAGAGAGAGATT 4412

RESULT 15
US-08-476-900A-4/C
: Sequence 4, Application US/08476900A
: Patent No. 6031150
: GENERAL INFORMATION:
: APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
: TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
: Patent No. 6031150
: TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
: TITLE OF INVENTION: Infancy
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:

```

RESULT 15
 US-08-476-900A-4/c
 ; Sequence 4, Application US/084769300A
 ; Patent No. 6031150
 ; GENERAL INFORMATION:
 ; APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
 ; TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonilyurea Receptor
 ; Patent No. 6031150
 ; TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
 ; TITLE OF INVENTION: Infancy
 ; NUMBER OF SEQUENCES: 49
 ; CORRESPONDENCE ADDRESS:
 ;

REFERENCE/DOCKET NUMBER: BYLR-0003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-404-531B-27

Query Match 9.2%; Score 29.8; DB 2; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
CY 118 GCGAAATGGCGATGTGATGCGCAGCAGGATCTTCCATTTTGACCGCATCAACGTTT 177
DB 4025 GCGAGCTTGGCGATGTGATGCGCAGCAGGATCTTCCATTTTGACCGCATCAACGTTT 3966
CY 178 CGCAGCATGCGATGATGATCTGAAAGACTCT 206
DB 3965 CGGAAAAAGCGCAGAGAGAGAGAGATTT 3937

RESULT 11
US-08-476-900A-27/c
Sequence 27, Application US/08476900A
Patent No. 6031150
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 6031150
TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
TITLE OF INVENTION: Infancy
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6031150 is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-476-900A-27

Query Match 9.2%; Score 29.8; DB 3; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
CY 118 GCGAAATGGCGATGTGATGCGCAGCAGGATCTTCCATTTTGACCGCATCAACGTTT 177
DB 4025 GCGAGCTTGGCGATGTGATGCGCAGCAGGATCTTCCATTTTGACCGCATCAACGTTT 3966
CY 178 CGCAGCATGCGATGATGATCTGAAAGACTCT 206
DB 3965 CGGAAAAAGCGCAGAGAGAGAGAGATTT 3937

RESULT 12
US-08-488-546A-27/c
Sequence 27, Application US/08488546A
Patent No. 6054313
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 6054313
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6054313 is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,546A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,531
FILING DATE: 15-MARCH-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533


```
QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 300
DB 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 300
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 301 TTGTGTGACTGTGTAATCTTTTC 324

RESULT 2
US-09-471-803A-1
; Sequence 1, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, WALTER
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnf
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brne
; OTHER INFORMATION: ATCC14752
US-09-471-803A-1

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCTGATATTCCTGTTGTGCGAGTATGTCAGTCACTTACTTT 60
DB 853 ATGACACTGATTTCTCTGATATTCCTGTTGTGCGAGTATGTCAGTCACTTACTTT 912
QY 61 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAAAATTTGGGGC 120
DB 913 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAAAATTTGGGGC 972
QY 121 AAATGGCGATGTGATGACGACGAGATCTTGACATTTTGAACCGCATCAAGTTTCCG 180
DB 973 AAATGGCGATGTGATGACGACGAGATCTTGACATTTTGAACCGCATCAAGTTTCCG 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTTGGTCTCAATGGCGGTTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTAACCCTTTGGTCTCAATGGCGGTTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 1152
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 1153 TTGTGTGACTGTGTAATCTTTTC 1176

RESULT 3
US-09-471-803A-6
```

```
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, WALTER
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnf
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brne
; OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCTGATATTCCTGTTGTGCGAGTATGTCAGTCACTTACTTT 60
DB 853 ATGACACTGATTTCTCTGATATTCCTGTTGTGCGAGTATGTCAGTCACTTACTTT 912
QY 61 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAAAATTTGGGGC 120
DB 913 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAAAATTTGGGGC 972
QY 121 AAATGGCGATGTGATGACGACGAGATCTTGACATTTTGAACCGCATCAAGTTTCCG 180
DB 973 AAATGGCGATGTGATGACGACGAGATCTTGACATTTTGAACCGCATCAAGTTTCCG 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTTGGTCTCAATGGCGGTTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTAACCCTTTGGTCTCAATGGCGGTTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGGCTTGGCGGTGACCATCGTT 1152
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 1153 TTGTGTGACTGTGTAATCTTTTC 1176

RESULT 4
US-09-252-991A-13729/c
; Sequence 13729, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
```

C 27	29.6	9.1	2454	3	US-08-488-546A-32	Sequence 32, App
------	------	-----	------	---	-------------------	------------------

C 28	29.6	9.1	4931.3	US-08-726-320-2	Sequence 2, Appl1
C 29	29.6	9.1	4931.1	US-09-208-716-2	Sequence 2, Appl1
C 30	29.4	9.1	61.1	US-09-976-594-152	Sequence 152, Appl1
C 31	29.4	9.1	12588	US-08-387-943C-1	Sequence 1, Appl1
C 32	29.2	9.0	867	US-09-216-339B-40	Sequence 340, Appl1
C 33	29.2	9.0	867	US-09-216-339B-42	Sequence 342, Appl1
C 34	29.2	9.0	1357	US-09-216-339B-343	Sequence 343, Appl1
C 35	29.2	9.0	1357	US-09-216-339B-345	Sequence 345, Appl1
C 36	29.2	9.0	9828	US-08-561-527-41	Sequence 41, Appl1
C 37	28.8	8.9	505	US-09-621-976-15639	Sequence 15639, Appl1
C 38	28.4	8.8	455	US-08-387-943C-155	Sequence 35, Appl1
C 39	28.4	8.8	7304	US-09-453-702B-174	Sequence 174, Appl1
C 40	28.2	8.7	387	US-08-651-155B-174	Sequence 74, Appl1
C 41	28.2	8.7	387	US-09-194-036B-174	Sequence 74, Appl1
C 42	28.2	8.7	5099	US-09-610-040-5	Sequence 5, Appl1
C 43	28.2	8.7	1230025	US-09-198-452A-1	Sequence 1, Appl1
C 44	28	8.6	531	US-08-671-548C-1	Sequence 7, Appl1
C 45	28	8.6	1756	US-09-205-258-173	Sequence 173, Appl1

ALIGNMENTS

RESULT 1
US-09-471-803A-4

Patent No. 6613545

APPLICANT: KENNERKNECHT, NICOLE
ADDRESSEE: SAM HERMAN

APPLICANT: EGGELING, LOTHAR

INVENTOR:	TITLE OF INVENTION:	TITLE OF INVENTION:	TITLE OF INVENTION:
1. INVENTOR:	2. TITLE OF INVENTION:	3. TITLE OF INVENTION:	4. TITLE OF INVENTION:
5. INVENTOR:	6. TITLE OF INVENTION:	7. TITLE OF INVENTION:	8. TITLE OF INVENTION:
9. INVENTOR:	10. TITLE OF INVENTION:	11. TITLE OF INVENTION:	12. TITLE OF INVENTION:
13. INVENTOR:	14. TITLE OF INVENTION:	15. TITLE OF INVENTION:	16. TITLE OF INVENTION:
17. INVENTOR:	18. TITLE OF INVENTION:	19. TITLE OF INVENTION:	20. TITLE OF INVENTION:
21. INVENTOR:	22. TITLE OF INVENTION:	23. TITLE OF INVENTION:	24. TITLE OF INVENTION:
25. INVENTOR:	26. TITLE OF INVENTION:	27. TITLE OF INVENTION:	28. TITLE OF INVENTION:
29. INVENTOR:	30. TITLE OF INVENTION:	31. TITLE OF INVENTION:	32. TITLE OF INVENTION:
33. INVENTOR:	34. TITLE OF INVENTION:	35. TITLE OF INVENTION:	36. TITLE OF INVENTION:
37. INVENTOR:	38. TITLE OF INVENTION:	39. TITLE OF INVENTION:	40. TITLE OF INVENTION:
41. INVENTOR:	42. TITLE OF INVENTION:	43. TITLE OF INVENTION:	44. TITLE OF INVENTION:
45. INVENTOR:	46. TITLE OF INVENTION:	47. TITLE OF INVENTION:	48. TITLE OF INVENTION:
49. INVENTOR:	50. TITLE OF INVENTION:	51. TITLE OF INVENTION:	52. TITLE OF INVENTION:
53. INVENTOR:	54. TITLE OF INVENTION:	55. TITLE OF INVENTION:	56. TITLE OF INVENTION:
57. INVENTOR:	58. TITLE OF INVENTION:	59. TITLE OF INVENTION:	60. TITLE OF INVENTION:
61. INVENTOR:	62. TITLE OF INVENTION:	63. TITLE OF INVENTION:	64. TITLE OF INVENTION:
65. INVENTOR:	66. TITLE OF INVENTION:	67. TITLE OF INVENTION:	68. TITLE OF INVENTION:
69. INVENTOR:	70. TITLE OF INVENTION:	71. TITLE OF INVENTION:	72. TITLE OF INVENTION:
73. INVENTOR:	74. TITLE OF INVENTION:	75. TITLE OF INVENTION:	76. TITLE OF INVENTION:
77. INVENTOR:	78. TITLE OF INVENTION:	79. TITLE OF INVENTION:	80. TITLE OF INVENTION:
81. INVENTOR:	82. TITLE OF INVENTION:	83. TITLE OF INVENTION:	84. TITLE OF INVENTION:
85. INVENTOR:	86. TITLE OF INVENTION:	87. TITLE OF INVENTION:	88. TITLE OF INVENTION:
89. INVENTOR:	90. TITLE OF INVENTION:	91. TITLE OF INVENTION:	92. TITLE OF INVENTION:
93. INVENTOR:	94. TITLE OF INVENTION:	95. TITLE OF INVENTION:	96. TITLE OF INVENTION:
97. INVENTOR:	98. TITLE OF INVENTION:	99. TITLE OF INVENTION:	100. TITLE OF INVENTION:

FILE REFERENCE: 21123/265496/MAS

CURRENT FILING DATE: 1999-12-23

PRIOR FILING DATE: 1999-10-27

SOFTWARE: PatentIn Ver. 2.1

; LENGTH: 324

ORGANISM: *Corynebacterium glutamicum*

[illegible]

OTHER INFORMATION: b7NE
OTHER INFORMATION: ATCC 14723

US-09-471-803A-4

Query Match	100.0%;	Score 324;	DB 4;	Length 324;
Best local Similarity	100.0%;	Pred. No. 2.4e-107;		

```
Matches 324; conservative 0; mismatches 0; indels 0; gaps 0;
```

A

22

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

181 AGGCAATGCGATAGATCTCGAAGGACTCTAAGCTTGGTCTCAATTGCGCGATTACAGTG 2

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at <http://wipo.int/pub/published-pct-sequences>. (Updated on 27-OCT-2003 to
 CC standardise OS field)

XX Sequence 470 BP; 122 A; 131 C; 102 G; 106 T; 0 U; 9 Other;

Query Match 9.9%; Score 32; DB 8; Length 470;

Best Local Similarity 51.9%; Pred. No. 2.2; Mismatches 0; Gaps 0;

Matches 68; Conservative 0; Indels 63; Indels 0; Gaps 0;

QY	181	AGCAATGCGATGATCTGAAGACTTACCTTGGTCTCATTCGCCGTTGCGATTACAGTG	240
DB	384	ATCAAGNACAGGAGNAGATAGGAGACCTTGAGNCCATCCAGTTGAGGATGCCAG	325
QY	241	GTGCGCGATCTTCTGGCGGTGACGCACTTGTAGGCTTGGCGCTGGCACCATCCTT	300
DB	324	GGTCACCAAGTTGCCACCCGCTGGGCAAGTTTCACCATTTGGCAAGGGCACCAAGCCT	265
QY	301	TTGTGTGACT	311
DB	264	TGGGTGAGCCT	254

Search completed: April 19, 2004, 11:42:03
 Job time : 153.886 secs

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published-pct-sequences](http://wipo.int/pub/published-pct-sequences). (Updated on 27-Oct-2003 to
 CC standardise OS field)

CC Sequence 376 BP; 97 A; 105 C; 82 G; 92 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 376;

Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 58; Indels 0; Gaps 0;

Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTTAACCTTTGCTCATTCGCCCTTGCATTACAGTGTGCG 246
 DB 375 GAGAGAGAGAGACATTAAGGAACTTTGAGACATCCAGTTGAGATGCCAGGCTCAC 316
 QY 247 CATCTTCTGGCGGTGACGACACCTTGTGAGCGTGGCGCTGGACCATCGTTTGT 306
 DB 315 CAGTTTCCACCCGCTGTGGGCAACGTTTCACTTGCACATTGGCAAGGCAACGCTTGGGTG 256
 QY 307 GGAAT 311
 DB 255 AGCCT 251

RESULT 14
 ACL20868/c
 ID ACL20868 standard; DNA; 392 BP.

AC 20868;
 AC 20868;
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10850.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KM gene; ss.

OS *Hordeum vulgare*; var. (cul.Akashiniki).

PN WO2003057877-A1.

PD 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

PR 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 27-SEP-2002; 2002JP-00327515.

PA (UYNI-) UNIV JAPAN OKAYAMA.

PI Sato K, Takeda K, Kohara Y;

WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.

PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at [ftp.wipo.int/pub/published-pct-sequences](http://wipo.int/pub/published-pct-sequences). (Updated on 27-Oct-2003 to
 CC standardise OS field)

CC Sequence 392 BP; 100 A; 105 C; 85 G; 102 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 392;

Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 58; Indels 0; Gaps 0;

Matches 67; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTTAACCTTTGCTCATTCGCCCTTGCATTACAGTGTGCG 246
 DB 391 GAGAGAGAGAGACATTAAGGAACTTTGAGACATCCAGTTGAGATGCCAGGCTCAC 332
 QY 247 CATCTTCTGGCGGTGACGACACCTTGTGAGCGTGGCGCTGGACCATCGTTTGT 306
 DB 331 CAGTTTCCACCCGCTGTGGGCAACGTTTCACTTGCACATTGGCAAGGCAACGCTTGGGTG 272
 QY 307 GGAAT 311
 DB 271 AGCCT 267

RESULT 15
 ACL20918/c
 ID ACL20918 standard; DNA; 470 BP.

AC 20918;
 AC 20918;
 DT 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10909.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KM gene; ss.

OS *Hordeum vulgare*; var. (cul.Haruna Nijo).

PN WO2003057877-A1.

PD 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

PR 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 27-SEP-2002; 2002JP-00327515.

PA (UYNI-) UNIV JAPAN OKAYAMA.

PI Sato K, Takeda K, Kohara Y;

WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.

PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX

Query Match	Similarity	Score	DB	Length
Best Local	58.6%	Pred. No. 5.4		
Matches	75	Conservative	0	Mismatches 51; Indels 2; Gaps 1;
Oy	93	GCCCCCTACGTAATCATCAATTGTGGGCAAAATGGCATGTGTGATGCCAGACGAATCT	152	
Db	5382	GCACACACGTGCATGTGCATCATGTGTGGAAAGAGGGCTGGGTGAGACATGACATCT	5323	
Oy	153	TGCCATTTTGACCGGATTAACGTTTGGAGCAATGCGATAGATCTGAAGACTTAACCTT	212	
Db	5322	TGTACACTGGC--CAACAGAAATATGGACAGANGTGGCTGGGTGTGTGTTCACACTG	5265	
Oy	213	TGCTCTCA 220		
Db	5264	TAAATCCA 5257		
RESULT 12				
ACF39600/c				
ID	ACF39600	standard; cDNA; 6672 BP.		
XX	AC	ACF39600;		
XX	DT	29-SEP-2003 (first entry)		
XX	DE	Human MHC class II transactivator encoding cDNA SEQ ID NO:3.		
XX	KM	Human; major histocompatibility complex class II transactivator;		
XX	KM	MHC class II transactivator; antisense modulation; immunosuppressive;		
XX	KM	antimicrobial; antidiabetic; antirheumatic; antiarthritic; cytostatic;		
XX	KM	nocrotropic; neuroprotective; immunostimulant; autoimmune disorder;		
XX	KM	MHC class II transactivator inhibitor; infection; transplant rejection;		
XX	KM	diabetes; rheumatoid arthritis; cancer; Alzheimer's disease;		
XX	KM	multiple sclerosis; severe combined immunodeficiency disease; gene; ss.		
XX	OS	Homo sapiens.		
XX	FH	Key	Location/Qualifiers	
XX	FT	CDS	139..3531	
XX	FT		/*tag= a	
XX	FT		/product= "MHC class II transactivator"	
XX	PN	MO2003050247-A2.		
XX	PD	19-JUN-2003.		
XX	PF	04-DEC-2002; 2002MO-US038616.		
XX	PR	05-DEC-2001; 2001US-00063366.		
XX	PA	(ISIS-) ISIS PHARM INC.		
XX	PI	Bennett FC, Dobie KW;		
XX	PI	WPI; 2003-577294/54.		
XX	DR	P-PSDB; ABR82108.		
XX	PT	New antisense oligonucleotides for modulating MHC class II transactivator		
XX	PT	gene expression, particularly useful for treating autoimmune disorders		
XX	PT	such as transplant rejection, Alzheimer's disease, or multiple sclerosis,		
XX	PT	or infection.		
XX	PS	Example 13; Page 93-99; 129pp; English.		
XX	XX	The present invention describes a compound (I) that is 8-50 nucleobases		
XX	XX	in length: (a) targets a nucleic acid molecule encoding major		
XX	XX	histocompatibility complex (MHC) class II transactivator, and		
XX	XX	specifically hybridises with the nucleic acid encoding the MHC class II		
XX	XX	transactivator; and inhibits the expression of MHC class II		
XX	XX	transactivator; or (b) specifically hybridises with at least an 8-		
XX	XX	nucleobase portion of an active site on a nucleic acid molecule encoding		
XX	XX	MHC class II transactivator. (I) has immunosuppressive, antimicrobial,		
XX	XX	antidiabetic, antirheumatic, antiarthritic, cytostatic, nocrotropic,		

Query Match	10.0%	Score 32.4	DB 8	Length 6672
Best Local Similarity	58.6%	Pred. No. 5.4	SI	Indels 2
Matches 75	Conservative 0	Mismatches 51	Gaps 1	
93	GGCCCTACGTGAATCACAAATTGTGGCGAATAATGGCGATGTGATGCCAGACGATCT	152		
Db	GCACACAGTGCATGTGCACATGTGTGGAAGAAGAGGGTGGGTGAGACAGTACGAACTCT	5322		
Qy	153	TGCCATTTTGACCCGATCAACGTTTCGAGAGAAAGCGATAGTCGGAAGACTCAACCT	212	
Db	5322	TGTCACTGCGC--CAACAGAAATATGCGAAGATGGGCTGGGTGTGTGCTTCAACCTG	5264	
Qy	213	TGTCTCA 220		
Db	5264	TATCCCA 5257		
RESULT 13				
ACTL20925/C				
ID	ACTL20925	standard; DNA; 376 BP.		
XX	ACL20925;			
XX	27-OCT-2003	(revised)		
DT	17-OCT-2003	(first entry)		
XX	DNA clone originating in barley containing SNP encoding sequence #10916.			
XX	Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;			
KW	gene; ss.			
OS	Hordeum vulgare; var. (cul.Akashinriki).			
XX	WO2003057877-A1.			
XX	17-JUL-2003.			
XX	16-DEC-2002; 2002WO-IB005403.			
XX	20-DEC-2001; 2001JP-00397059.			
XX	20-DEC-2001; 2001JP-00387131.			
PR	20-DEC-2001; 2001JP-00403299.			
PR	20-DEC-2001; 2001JP-00403300.			
PR	27-SEP-2002; 2002JP-00327515.			
XX	(UTNT-) UNIV JAPAN OKAYAMA.			
XX	Sato K, Takeda K, Kohara Y;			
XX	WPI; 2003-587127/55.			
XX	Single nucleotide polymorphism sites in barley varieties and DNA			
PT	sequences containing them for analysis and identification of barley			
PT	varieties and production of barley transformants with desired			
XX	characteristics.			
XX	Disclosure; SEQ ID XX; 284pp; Japanese.			
XX				

XX The invention relates to an isolated polynucleotide (ACF6435-ACF64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridization. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present invention represents a P. acnes DNA contig which is specifically claimed in the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 66788 BP; 13534 A; 21369 C; 19347 G; 12530 T; 0 U; 8 Other;

XX Query Match 10.1%; Score 32.6; DB 7; Length 66788; Best Local Similarity 52.6%; Pred. No. 13; Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

XX QY 154 GCCATTGACCGCATCAACGTTGGACGATCGATGATGAAAGCTTAACCTT 213
DB 40085 GTCATTCTGATCCGCGGACGAAACCCCGAGCTTCTTGACCCGCTGCTACCTCT 40026

XX QY 214 GGTCTCATGCGCGTTGGATTACAGTGTGGCGCATCTTCTGGCGTACGACGACTTG 273
DB 40025 GGGCGTTTTCACGCTGATGCGATGAGAGCTCCGATCTCATGCGCGTCTCAAGATT 39966

XX QY 274 TTGAGCGTTGCGCT 288
DB 39965 CTCAGGTTCCGCT 39951

XX AC ABL68121;

XX DT 15-MAY-2002 (first entry)

XX DE Ovary cancer related gene sequence SEQ ID NO:6458.

XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid; stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous; K120; gene; ds.

XX OS Homo sapiens.

XX PN WO200194629-A2.

XX PD 13-DEC-2001.

XX PP 30-MAY-2001; 2001WO-US010838.

XX PR 05-JUN-2000; 2000US-0209473P.

XX PR 05-JUN-2000; 2000US-0209531P.

XX PR 18-SEP-2000; 2000US-0233133P.

PR 18-SEP-2000; 2000US-0233617P.

PR 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234034P.

PR 20-SEP-2000; 2000US-0234052P.

PR 22-SEP-2000; 2000US-0234509P.

PR 22-SEP-2000; 2000US-0234567P.

PR 25-SEP-2000; 2000US-0234923P.

PR 25-SEP-2000; 2000US-0234924P.

PR 25-SEP-2000; 2000US-0235077P.

PR 25-SEP-2000; 2000US-0235082P.

PR 25-SEP-2000; 2000US-0235134P.

PR 25-SEP-2000; 2000US-0235280P.

PR 26-SEP-2000; 2000US-0235637P.

PR 26-SEP-2000; 2000US-0235638P.

PR 27-SEP-2000; 2000US-0235711P.

PR 27-SEP-2000; 2000US-0235720P.

PR 27-SEP-2000; 2000US-0235840P.

PR 27-SEP-2000; 2000US-0235863P.

PR 28-SEP-2000; 2000US-0236028P.

PR 28-SEP-2000; 2000US-0236032P.

PR 28-SEP-2000; 2000US-0236033P.

PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.

PR 28-SEP-2000; 2000US-0236111P.

PR 29-SEP-2000; 2000US-0236842P.

PR 29-SEP-2000; 2000US-0236891P.

PR 02-OCT-2000; 2000US-0237112P.

PR 02-OCT-2000; 2000US-0237173P.

PR 02-OCT-2000; 2000US-0237278P.

PR 02-OCT-2000; 2000US-0237284P.

PR 02-OCT-2000; 2000US-0237295P.

PR 02-OCT-2000; 2000US-0237316P.

PR 03-OCT-2000; 2000US-0237425P.

PR 03-OCT-2000; 2000US-0237598P.

PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237606P.

PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.

PA (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S; PI Soppet DR, Weaver Z; DR WIPI 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 6458; 44bp; English.

XX The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophagus, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms tumor

XX SQ Sequence 6672 BP; 1392 A; 1868 C; 1458 T; 0 U; 0 Other;

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB372072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3144 BP, 1056 A, 676 C, 680 G, 732 T, 0 U, 0 Other;
 SQ Query Match 10.1%; Score 32.8; DB 4; Length 3144;
 Best Local Similarity 52.1%; Pred. No. 2.8;
 Matches 73; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 169 TCACGTTTCGACGCAATGCGATGATCTGAGAGCTTACTTGGTCTGACCTGCT 228
 DB 1512 TCTAAATTTCTAT 1453
 QY 229 GCGATTACAGTGGTGGCGCATCTTCTGGCGGTGACGACCTTGTGAGCGTTGCGCT 288
 DB 1452 TCGTTTGGGTTGCTGCTTCTTCTTTCGTCGCGGACGCAATCGTTGCTGCGTGA 1393
 QY 289 GGCACCATGCTTTTGTGTG 308
 DB 1392 CACAGCTTATGTTGTGTG 1373

RESULT 9
 AAS59515/C
 ID AAS59515 standard; DNA, 66788 BP.

XX AAS59515;
 AC
 XX
 DT 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #10.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacharya A;
 PI L'Almeida J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Claim 1; SEQ ID NO 10; 1069bp; English.

XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by *P. acnes*. The disorders include

CC SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and
 CC osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in
 CC infections of bone, joints and the central nervous system, however it is
 CC particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of *P. acnes* in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for *P. acnes* proteins. These antibodies can be
 CC used to downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU1313-42019 and AAU67473. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 66788 BP; 13534 A, 21369 C, 19347 G, 12530 T, 0 U, 8 Other;

QY Query Match 10.1%; Score 32.6; DB 4; Length 66788;
 Best Local Similarity 52.6%; Pred. No. 13;
 Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 154 GCCATTTTGACCGCATGATGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 213
 DB 40085 GTCAATTCGATCGCGGACGACGACGACGACGACGACGACGACGACGACGACGAC 40026

QY 214 GGTCTCATGCGCGTGGGATTAAGTGTGCGCGATCTTTTGGCGGTGACGACCTTG 273
 DB 40025 GGGCGTTTCGACCGCTGACGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 39966

QY 274 TTGAGCGTGGCGCT 288
 DB 39965 CTCAGGTCACGCT 39951

RESULT 10
 ACF64444/C
 ID ACF64444 standard; DNA, 66788 BP.

XX ACF64444;
 AC

XX 17-OCT-2003 (first entry)

XX Propionibacterium acnes DNA contig sequence #10.

XX Acne vulgaris; anti-seborrheic; dermatological; antibacterial;
 XX immunostimulant; immune response; vaccine; ds.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAM, Persing DH, Bhattacharya A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MT, Benson DR, Jones R, Carter D;

XX Barth B, Vallée-Douglas U;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a *P. acnes* protein.

XX Claim 1; SEQ ID NO 10; 1481bp; English.

CC Corynebacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
CC

SQ Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;

Query Match	Score	DB 5;	Length
100.0%;	324;	DB 5;	349980;
100.0%;	Prod No 3	10-89.	

Best Local Similarity	100.00%	File: NO. 2.15-55
Matches 324;	Conservative 0;	Mismatches 0;
		Indels 0;
		Gaps 0;

QY	1	ATGACAACGTGATTTCTCCCGTATCTCCCTGTGTGGCAGATGTGAGTCATTAATT	60
Db	277581	ATGACAACGTGATTTCTCCCGTATCTCCCTGTGTGTGGCAGATGTGAGTCATTAATT	277640
QY	61	GCCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATGACATTTGGGGC	120
Db	277641	GCCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATGACATTTGTGGGC	277700
QY	121	AAATGGCGATGTGGATGCGACGAGAACTCTTGCATTTGACCGCATCAACGTTTCGC	180
Db	277701	AAATGGCGATGTGGATGCGACGAGAACTCTTGCATTTGACCGCATCAACGTTTCGC	277760
QY	181	AGCAATCCGATTAATCTGAAGAATCTTAACCTTGTCTCATTTGCCGTGGCATTAACGTG	240
Db	277761	AGCAATCCGATTAATCTGAAGAATCTTAACCTTGTCTCATTTGCCGTGGCATTAACGTG	277820
QY	241	GTGGCGCATCTTTGTGGCGGTGCACGCAACTTGTGAAGCGTGGCGTGGCACCATCGTT	300
Db	277821	GTGGCGCATCTTTGTGGCGGTGCACGCAACTTGTGAAGCGTGGCGTGGCACCATCGTT	277880
QY	301	TTTGTTGGACTGGTGAAATCTTTTC	324
Db	277881	TTTGTTGGACTGGTGAAATCTTTTC	277904

RESULT 7
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

DT 20-NOV-2003 (first entry)

DE Rice gene, SEQ ID 5263.

KM plant; bacterial infection; fungal infection; viral infection; rice
KM gene; ds.

OS *Oryza sativa*.

PN WO2003000898-A1.

PD 03-JAN-2003

PF 22-JUN-2001; 2001WO-IB001105.

PR 22-JUN-2001; 2001WO-IB001105.

PA (SYGN) SYNGENTA PARTICIPATIONS AG
 XY

PI Chang H, Chen W, Cooper B, Glazebrook J, Gott SA, Hou Y;
PI Katsirji F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.

AS Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.
yy

CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC the expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match	Score	DB	Length
10.7%;	34.8;	7;	2000;
12.9%;	34.8;	7;	2000;

Matches	30;	Conservative	105;	Mismatches	97;	Indels	0;	Gaps	0
---------	-----	--------------	------	------------	-----	--------	----	------	---

QY 4 ACAACGATATTTCCCGGATCTCTCGTTGTGGAGAGATCTGGAGCATTAATTGGG 63

Db 901 AYAAMRRRTYTKMSRMYTMTMTKMTWMTCMQAMAYVATGMATMMWRMYNTYCY 842

QY 64 CTCGGGGCGGTCGTTCTTAAATCCTTAAGCCCTACGTGAATCACAATTTGTGGCGAA 123

Db 841 AMTCAKCKYKAMTKMTWTTMACAMBATSRWRALMAGMRKEYSKKRKAAYWRWRMRWRAGW 782

QY 124 ATGGCGATGTGGATGCCAGACGAATCCTTGCATTTTGGACCGCATCAAGTTTGGAGC 183

Db 781 ARWMSKSYFMKKKYATRYRKKMMAMTWMSRFRKSRFMSGGRMSWSAMRCSWMC 722

QY 184 AATGCGATATCTGAAGAATCTAAACCTTGGTCTCATTCGCGTTGGCATTA 235

Db 721 AKTKIASAMWTKRASKSTRYRRMRMKKMTYTYRFRMRSCMTPRAMSR 670

RESULT 8
ABL27130/c
ID ABL27130 standard; DNA; 3144 BP.

AC ABL27130;

DT 26-MAR-2002 (first entry)

Drosophila melanogaster genomic polynucleotide SEQ ID NO 32863.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

Drosophila melanogaster

PN WO200171042-A2

PD 27-SEP-2001.
xx

PF 23-MAR-2001; 2001WO-US009231.
YY

PR 23-MAR-2000; 2000US-019163/E.
PR 11-JUL-2000; 2000US-00614150-1

XX
PA
(PEKE) PE CORP NY.

xx Venter JC, Adams N
PI

XX
DR WPT: 2001-656860/75.

PT New isolated nucleic

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.

PS Claim 1; SEQ ID NO 32863; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

Tue Apr 20 06:47:19 2004

US-10-608-504-4.ing

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 150.886 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-4

Sequence: 324
1 atgacaactgattctctctg.....ttgactgtgacatcttctc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq.29qand04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2003as:*\n8: geneseqn2003bs:*\n9: geneseqn2003cs:*\n10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	324	100.0	324	AAH21111	Aah21111 C. glutam
2	324	100.0	324	AAH65255	Aah65255 C. glutam
3	324	100.0	324	ACA01970	ACA01970 C. glutam
4	324	100.0	1271	AAH21109	Aah21109 C. glutam
5	324	100.0	1271	AAH21112	Aah21112 C. glutam
6	324	100.0	349980	AAH64966	Aah64966 C. glutam
7	34.8	10.1	2000	ADA71938	Ada71938 Rice gene
8	32.8	10.1	3144	ABL27130	AbL27130 Drosophila
9	32.6	10.1	66788	AAAS9515	Aaas9515 Propionib
10	32.6	10.1	66788	ACF64444	Acf64444 Propionib
11	32.4	10.0	66782	AB168121	Ab168121 Ovary can
12	32.4	10.0	66772	ACF38600	Acf38600 Human MHC
13	32.2	9.9	376	ACL20925	ACL20925 DNA clone
14	32.2	9.9	392	ACL20868	ACL20868 DNA clone
15	32	9.9	470	ACL20918	ACL20918 DNA clone
16	32	9.9	476	ACL20933	ACL20933 Human CDN
17	32	9.9	820	AAH05013	Aah05013 Human CDN
18	32	9.9	1808	AAH18459	Aah18459 Human CDN
19	31.8	9.8	2091	ADB05513	Adb05513 Alloiococ
20	31.8	9.8	2091	ADB06509	Adb06509 Alloiococ
21	31.8	9.8	2091	ADB06511	Adb06511 Alloiococ
22	31.8	9.8	110000	ADB12064_00	ADB12064 Alloiococ
23	31.8	9.8	110000	ADB12064_01	Continuation (2 of

24	31.6	9.8	1404	8	ADA30052	Ada30052 DNA encod
25	31.4	9.7	2387	2	AAV52189	Aav52189 Streptococ
26	31.4	9.7	110000	5	ABBS6454_07	Continuation (8 of
27	31	9.6	4590	5	AAH24065	Aah24065 Yeast AOD
28	31	9.6	6415	2	AAH13212	Aah13212 Enterococ
29	31	9.6	6415	6	ABBS9007	Abbs9007 Enterococ
30	30.8	9.5	5949	2	AAH30053	Aah30053 S. pneumo
31	30.8	9.5	265118	5	AAH41227	Aah41227 Pyrococcu
32	30.6	9.4	413	7	ACL20863	ACL20863 DNA clone
33	30.6	9.4	789	7	ACA50939	Acas50939 Prokaryot
34	30.4	9.4	386	8	ACL20939	ACL20939 DNA clone
35	30.4	9.4	4739	6	ABX03452	Abx03452 Negative-
36	30.2	9.3	373	8	ACL20861	ACL20861 DNA clone
37	30.2	9.3	448	8	ACL20929	ACL20929 DNA clone
38	30.2	9.3	505	8	ACL20950	ACL20950 DNA clone
39	30.2	9.3	515	8	ACL20954	ACL20954 DNA clone
40	30.2	9.3	528	8	ACL20920	ACL20920 DNA clone
41	30.2	9.3	529	8	ACL20912	ACL20912 DNA clone
42	30.2	9.3	533	8	ACL20906	ACL20906 DNA clone
43	30.2	9.3	537	8	ACL20947	ACL20947 DNA clone
44	30.2	9.3	538	8	ACL20907	ACL20907 DNA clone
45	30.2	9.3	540	8	ACL20924	ACL20924 DNA clone

ALIGNMENTS

RESULT 1
AAH21111
ID AAH21111 standard; DNA; 324 BP.
XX AAH21111;
AC AAH21111;
DT 05-SEP-2001 (first entry)
XX C. glutamicum brne DNA.
DE
KW L-amino acid production; brnF, brnE; branched-chain amino acid;
KW Corynebacterium bacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.
XX
CS Corynebacterium glutamicum.
XX
FN EP1096010-Al.
XX
PD 02-MAY-2001.
XX
PF 11-OCT-2000; 2000EP-00122057.
XX
PR 27-OCT-1999; 99DE-01051708.
XX
PA (DEGS) DEGUSA AG.
PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
DR Kennertnecht N, Eggeling L, Sahn H, Pfeffertle W;
XX WPI, 2001-391595/42.
XX F-PSDB; AAB86248.
XX
PT New export genes from corynebacterium bacteria, useful for increasing
PT fermentative production of branched-chain amino acids.
XX
PS Claim 5; Page 16; 23pp; German.
XX
CC This invention describes a novel isolated polynucleotide (I) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i) - (iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC corynebacterium microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of

```

/product="putative mannosyltransferase A"
/protein_id="AA064208.1"
/db_xref="GI:28974207"
/translation="1HDLIPITSPQFPISCELLQWIERIKIDPVLATISQKTN
DLKYEKFKIKIPLGSEFTELKVRHIEKRFELAVGTIERAKAKAIVNTFLQD
EIERDLITVKGIGVDEDFEKECKKHNIKXIKFVSDRLGSLYSTETILASVD
EGYGLPMVNASIYGACITARDIPVFRDLIGENATYFKDIDISLYETIKVYSKIKNRRI
GITWEYAREILIKIKEDL"
gene
711..1856
/gene="mtfb"
711..1856
/gene="mtfb"
/feature="Cf0002"
/codon_start=1
/translation="putative mannosyltransferase B"
/protein_id="AA064209.1"
/db_xref="GI:28974208"
/translation="MINILIDTRPLASNLGTGRTYVEIAKLNKNEFKPYDYGFIL
DELYGDNQOKKPSIKSHLINIKRILDKPTIKKVKQKLIKSLREKYSINIL
YQENRPIELKAKTITVTTHILAPMSYEPHPEKRIIDPEONFIPNLSKTDHITVS
NAIKHGIKKNLISADKISVITNGIDENIRPKRQKOTINIKDLEINPILFVGS1
EPRKQITLIIQAVNELNINIDLVIGAGWENSISHLIQNHEIKFLGFTPPDDLA
LRYSSATIFVPSIYEGFIPLEIAACAPILSDIEVFREIYGNVAFESPLNAKE
LAEKLNINNSDKLSIMKNGILISKYSWKCKIQEHKELIKILD"
2182..5010
/gene="sapA"
2182..5010
/gene="sapA"
/feature="Cf0003; similar to sap"
/codon_start=1
/translation="surface array protein"
/protein_id="AA064210.1"
/db_xref="GI:28974209"
/translation="MLNKTQVSMLYITTINGASSEDGKRYMLDYANNNTLGVSLANI
MLDSGNAKFFGSLIAGNKDPTKIVISIALGNISDVQINVTAKITGGSEPTDSK
GNIVSVASLKGDLIGAMINSMVNGSASKAIPEAKAASVYFADATLQKISLDE
GYTSKLISEINSSASDIDKYSEIDALSKSLPNPSPGYDTEGNDMLKGTDLDTFNGE
TIVGNGTNSKSLAPFKIDGAGADTLNQVLANNSASAPAKSEDLVYALGVNVE
SVNIVDLDPDRAITLNGSGTGAMNALLDGSKLAGLTDISGKLNIVNGSVALT
AARAKSVNDVAVATITLTPKRAVNVKVAAGDVTLDTEASASISVDANGKVTLT
KATVENSIVKATINUTITRAGDKNTVTLBNVTLDSSAAAIIDPKATILANNVY
KMDGSDSKIAHKSSAETLNLNFSKKEATFALVNOVKVATITNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELVADGLDVTSIDASGENKATVFGSNG
DALAGKVSAGANDTLVLESNKLSAGSVIDGAGADTIVLKSALADATLGMKNI
EAVOVDKLTANTNISTAPFTIIVAGDETSOTKLINDOKIDLAATYLAASKITM
DLNEKSGDNDISNLVNAKTADNNVTVNGADKQKIVIEGLSVITISVADSTYAN
SIDKPTDYKSIILINGDAGITLKGFTVADYKRVSMIDASAKAGLTFPASAITANG
ATTGSGADSVITVKGALANLTGEGSDTYTLKGGTSRDYITINDRSDDKIVLDGT
TFASAKAIEKTLANTLTFSDYVNOASGNGNTPIVKIFHYONDYIIVYDGAGATP
TKADIVIKLAGIHEITGTNTITADSQ"
gene
complement(5749..9069)
/gene="sapA5"
complement(5749..9069)
/feature="Cf0004; similar to sap"
/codon_start=1
/translation="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTQVSMLYITTINGASSEDGKRYMLDYANNNTLGVSLANI
MLDSGNAKFFGSLIAGNKDPTKIVISIALGNISDVQINVTAKITGGSEPTDSK
GNIVSVASLKGDLIGAMINSMVNGSASKAIPEAKAASVYFADATLQKISLDE
GYTSKLISEINSSASDIDKYSEIDALSKSLPNPSPGYDTEGNDMLKGTDLDTFNGE
TIVGNGTNSKSLAPFKIDGAGADTLNQVLANNSASAPAKSEDLVYALGVNVE
SVNIVDLDPDRAITLNGSGTGAMNALLDGSKLAGLTDISGKLNIVNGSVALT
AARAKSVNDVAVATITLTPKRAVNVKVAAGDVTLDTEASASISVDANGKVTLT
KATVENSIVKATINUTITRAGDKNTVTLBNVTLDSSAAAIIDPKATILANNVY
KMDGSDSKIAHKSSAETLNLNFSKKEATFALVNOVKVATITNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELVADGLDVTSIDASGENKATVFGSNG
DALAGKVSAGANDTLVLESNKLSAGSVIDGAGADTIVLKSALADATLGMKNI
EAVOVDKLTANTNISTAPFTIIVAGDETSOTKLINDOKIDLAATYLAASKITM
DLNEKSGDNDISNLVNAKTADNNVTVNGADKQKIVIEGLSVITISVADSTYAN
SIDKPTDYKSIILINGDAGITLKGFTVADYKRVSMIDASAKAGLTFPASAITANG
ATTGSGADSVITVKGALANLTGEGSDTYTLKGGTSRDYITINDRSDDKIVLDGT
TFASAKAIEKTLANTLTFSDYVNOASGNGNTPIVKIFHYONDYIIVYDGAGATP
TKADIVIKLAGIHEITGTNTITADSQ"
gene
complement(5749..9069)
/gene="sapA5"
complement(5749..9069)
/feature="Cf0004; similar to sap"
/codon_start=1
/translation="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTQVSMLYITTINGASSEDGKRYMLDYANNNTLGVSLANI
MLDSGNAKFFGSLIAGNKDPTKIVISIALGNISDVQINVTAKITGGSEPTDSK
GNIVSVASLKGDLIGAMINSMVNGSASKAIPEAKAASVYFADATLQKISLDE
GYTSKLISEINSSASDIDKYSEIDALSKSLPNPSPGYDTEGNDMLKGTDLDTFNGE
TIVGNGTNSKSLAPFKIDGAGADTLNQVLANNSASAPAKSEDLVYALGVNVE
SVNIVDLDPDRAITLNGSGTGAMNALLDGSKLAGLTDISGKLNIVNGSVALT
AARAKSVNDVAVATITLTPKRAVNVKVAAGDVTLDTEASASISVDANGKVTLT
KATVENSIVKATINUTITRAGDKNTVTLBNVTLDSSAAAIIDPKATILANNVY
KMDGSDSKIAHKSSAETLNLNFSKKEATFALVNOVKVATITNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELVADGLDVTSIDASGENKATVFGSNG
DALAGKVSAGANDTLVLESNKLSAGSVIDGAGADTIVLKSALADATLGMKNI
EAVOVDKLTANTNISTAPFTIIVAGDETSOTKLINDOKIDLAATYLAASKITM
DLNEKSGDNDISNLVNAKTADNNVTVNGADKQKIVIEGLSVITISVADSTYAN
SIDKPTDYKSIILINGDAGITLKGFTVADYKRVSMIDASAKAGLTFPASAITANG
ATTGSGADSVITVKGALANLTGEGSDTYTLKGGTSRDYITINDRSDDKIVLDGT
TFASAKAIEKTLANTLTFSDYVNOASGNGNTPIVKIFHYONDYIIVYDGAGATP
TKADIVIKLAGIHEITGTNTITADSQ"

```

```

NLVNTGAFSIAATATSLKLNINLSINGVADLNSVNYGTATLASLEANINVGEFKL
GTTTAKGVDPEINENVGALLTGATITSNGASVSISSATGVTLGAVATQGNLTINA
GNTLGNITITGAKGDIIVAGVTLGGVGTGTINSNAPNARSITSNBEVTVGSEISKNVEIT
AAAGGIDLNQVITGGAAADDAITLTKGDDTOTITASGDLSGGTLTLTLEATPKLSLD
ISAGVLSKATVITIDIKNVSVENKLVIVDQSDAETITANSATITLITSGDLGGG
ANTVAVAPDAVAATITIDLSGSLATGTLGDTITNNAQTLITTKSAGNDITITG
INADITVAGAGNDVFNVTAKIVTANTPEHATITDPSAGSIFPAASVAYKRSIV
DLSGRADLSAIAVLTDSDEATVYGYFTYNESLYLVAVATTATTAANDVLYLGTG
TIVDLSLITVNTDIVEA"
complement(9374..10762)
/codon_start=1
/evidence="not_experimental"
/translation="hypothetical protein Cf0005"
/protein_id="AA064212.1"
/db_xref="GI:28974211"
/translation="MONITINONKMSSEPYKLYRQWPKLINSIKLRHIMVLOGIR
RSGKSTVFKLLINHLKSDSKELIYINIDPEPSTKSKADVIDINDIYLSORLTKR
VKYLFDELSAVNWKYKSAVDSLPKRTITTSNSALNTQYATLISGRYLSROI
YPLSFEIIRKINNINDYFELIKENHVLNVQDMAKFGSVFVTSNENFEKRDITISY
YDAILKDCVANNQIIRDIAGFRELGFSLNFGALYSVNSLNAVKLNSEVQETIGY
LQCYLDCDELKMFYSILKEQMNKKLYLDCNGFLSLAFSDNDKGRALENVFCELY
KQGFELFPYKINDSECFIARKDKSIALQVCEINENKIKREIAGLNLPGVENEK
LITPDRWRETELQFPLRLGLNLDVYKRPFGGDNSEETREKSEKELKIMPFYEF
FSEFNETNIAL"
complement(10992..14765)
/gene="sapA6"
complement(10992..14765)
/gene="sapA6"
/feature="Cf0006; similar to sap"
/codon_start=1
/translation="surface array protein"
/protein_id="AA064213.1"
/db_xref="GI:28974212"
Query Match 11.8%; Score 38.2; DB 1; Length 65865;
Best Local Similarity 47.0%; Pred. No. 2.4;
Matches 118; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 5 CAACTGATTTCTCCTGATTCCTCTGTGTGCGAGATGTGCACTATTCTTGGCC 64
DB 45075 CAATCCTGTTGCGCTATGTTATATTTTGTGCTGATTCGAGATGTATGTACGC 45016
QY 65 TCCGGCGGTCCTGCTTCAATCCTTAAGCCCTAGCGATCAATTTGGGCAAAA 124
DB 45015 TTTAGCTGCTCCGCTGTTTAAATTCACACTAGACAGCTTTGCACTGGTTAATG 44956
QY 125 TGGCAGTGTGATCGCAGCAGAAATCCTTGCATTTTGACCGATCAAGCTTTGGCAGCA 184
DB 44955 TGATAGTGTGTTTGACGATTAATTAATTAAGCTTGTGTCGGAATCGGCTTAAGGCGAG 44896
QY 185 ATGCATATGATGTAGAGCTTAACCTTGGCTCATTCGCGTTCGAGTACAGTGTG 244
DB 44895 AGATTTTACCGGCGCAACATCGACTTTCCTGTTCCGCTTCGCTGAACGATTAATCG 44836
QY 245 CGCATCTTCTT 255
DB 44835 TTCCTGTGTT 44825
Search completed: April 19, 2004, 18:30:01
Job time : 1377.24 secs

```

DEFINITION	Sequence 1 from Patent EP1108790.
ACCESSION	AX120085 AX114121
VERSION	AX120085.1 GI:14036800
KEYWORDS	
SOURCE	
ORGANISM	Corynebacterium glutamicum
REFERENCE	Corynebacterium glutamicum
AUTHORS	Corynebacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
TITLE	1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
JOURNAL	Novel polynucleotides
FEATURES	Patent: EP 1108790-A 1 20-JUN-2001;
SOURCE	KYOWA HAKKO KOGYO CO., LTD. (JP)
ORIGIN	Location/Qualifiers
Query Match	100.0%; Score 324; DB 6; Length 349980;
Best Local Similarity	100.0%; Pred. No. 1,1e-88;
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AAGACACGATTTCCCTCGATATCCCTGTTGTGCGAGTATGACGATCTTTT 60
DB	277581 AAGACACGATTTCCCTCGATATCCCTGTTGTGCGAGTATGACGATCTTTT 277640
QY	61 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTAATCACATTTGTGGC 120
DB	277641 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTAATCACATTTGTGGC 277700
QY	121 AAAATGGGATGTGATGTCGACGAGGAATCCTTGATTTGACCGCATCAAGTTTCG 180
DB	277701 AAAATGGGATGTGATGTCGACGAGGAATCCTTGATTTGACCGCATCAAGTTTCG 277760
QY	181 ACGAATGCGATGATGATGTAAGACTCTAACCTTTGTTCAATGCGGTGGATTAAGTG 240
DB	277761 ACGAATGCGATGATGATGTAAGACTCTAACCTTTGTTCAATGCGGTGGATTAAGTG 277820
QY	241 GTGGCGCATCTTCTGGCGGTGAGGACCTTGATGAGGTTGCGCGTGGACCATCGTT 300
DB	277821 GTGGCGCATCTTCTGGCGGTGAGGACCTTGATGAGGTTGCGCGTGGACCATCGTT 277880
QY	301 TTGTGTGAATGCTGTAATCTTTTC 324
DB	277881 TTGTGTGAATGCTGTAATCTTTTC 277940

RESULT 15

AX121269/c

LOCUS AY211269 65865 bp DNA linear BCT 21-APR-2003

DEFINITION Campylobacter fetus strain 23D sap gene locus, partial sequence.

ACCESSION AY211269 AF027405 J05577 L15800 S44580 S76860

KEYWORDS AY211269.1 GI:28974206

SOURCE

ORGANISM Campylobacter fetus

Campylobacter fetus

Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.

1 (bases 1 to 65865)

Blaser, M.J. and Gottschlich, E.C.

Surface array protein of Campylobacter fetus. Cloning and gene structure

JOURNAL J. Biol. Chem. 265 (24), 14529-14535 (1990)

MEDLINE 90354448

PubMed 2387868

2 (bases 1 to 65865)

Blaser, M.J. and Gottschlich, E.C.

Surface array protein of Campylobacter fetus. Cloning and gene

REFERENCE

AUTHORS

TITLE

AUTHORS

TITLE

```

JOURNAL J. Biol. Chem. 265 (31), 19372-19390)
MEDLINE 91035477
PUBMED 2229082
REFERENCE 3 (bases 1 to 65865)
AUTHORS Tummuru,M.K. and Blaser,M.J.
TITLE Characterization of the Campylobacter fetus sapa promoter: evidence
that the sapa promoter is deleted in spontaneous mutant strains
JOURNAL 92394895
MEDLINE 1522068
PUBMED 1522068
REFERENCE 4 (bases 1 to 65865)
AUTHORS Tummuru,M.K. and Blaser,M.J.
TITLE Rearrangement of sapa homologs with conserved and variable regions
in Campylobacter fetus
JOURNAL 93348254
MEDLINE 8346244
PUBMED 8346244
REFERENCE 5 (bases 1 to 65865)
AUTHORS Dworkin,J., Tummuru,M.K. and Blaser,M.J.
TITLE A lipopolysaccharide-binding domain of the Campylobacter fetus
S-layer protein resides within the conserved N terminus of a family
of silent and divergent homologs
JOURNAL 95204338
MEDLINE 177 (7), 1734-1741 (1995)
PUBMED 177 (7), 1734-1741 (1995)
REFERENCE 6 (bases 1 to 65865)
AUTHORS Thompson,S.A., Shedd,O.L., Ray,K.C., Beins,M.H., Jorgensen,J.P. and
Blaser,M.J.
TITLE Campylobacter fetus surface layer proteins are transported by a
type I secretion system
JOURNAL 99069317
MEDLINE 180 (24), 6450-6458 (1998)
PUBMED 180 (24), 6450-6458 (1998)
REFERENCE 7 (bases 1 to 65865)
AUTHORS Tu,Z.-C., Wassenaar,T.M., Thompson,S.A. and Blaser,M.J.
TITLE Structure and genotypic plasticity of the Campylobacter fetus sap
locus
JOURNAL Mol. Microbiol. 48 (3), 685-698 (2003)
MEDLINE 12694614
PUBMED 12694614
REFERENCE 8 (bases 1 to 65865)
AUTHORS Blaser,M.J.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1993) Division of Infectious Diseases, Vanderbilt
University School of Medicine, Nashville, TN 37232, USA
MEDLINE 9 (bases 1 to 65865)
PUBMED 9 (bases 1 to 65865)
REFERENCE 9 (bases 1 to 65865)
AUTHORS Thompson,S.A., Shedd,O.L., Ray,K.C., Beins,M.H., Jorgensen,J.P. and
Blaser,M.J.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2003) Microbiology and Medicine, New York
University School of Medicine, YAMC ROOM 6006W 423 E. 23rd Street,
New York, NY 10016, USA
COMMENT On or before Mar 15, 2003 this sequence version replaced gi:255285,
gi:4009445, gi:144211, gi:128933, gi:913763.
FEATURES
source Location/Qualifiers
1..65865
/oranism="Campylobacter fetus"
/mol_type="genomic DNA"
/stain="2D"
/serotype="A"
/db_xref="taxon:196"
<1..714
/gene="mtfA"
<1..714
/gene="mtfA"
<1..714
/gene="mtfA"
/note="CF0001"
/codon_start=1
/transl_table=1

```

gene 3585..4769
/gene="Cg10004"
3585..4769
CDS /note="TIGR00611:recf: Recf protein"
/codon_start=1
/transl_table=11
/product="Recombinational DNA repair: ATPase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translation="WHIRSLLELDYRSWPELKVDLEPGITVFICNGNEKNTIVAGI
YLAHLSHRVSDAPLFAVRAENARVSAVAVQGRBLAHLIKPHAAQASLNTKV
TRRELIGVAVTLPFAPEDLALVGPSPARRRRLDIIATROPMAAGVADYKVLK
RNALLKTAATLALRRGYTEGGAALSLDITWDQGLARLGAEVMAAFALINELGRTY
EAYTTAPESRPAAVNYKTTIDQLSOFSEPDGITEATLTELAKROREIERGSS
VGRHDDVDIMLGDOPAKGPSHGETMSPLSIRIAEPNLKSGDIDPILIDDVSE
LDAGREKLVGIAQVEVQVLTAAVHDDPENLKVTLNQHTVTVQDITGRISILDV
Qp"
gene 4766..5302
/gene="Cg10005"
4766..5302
CDS /gene="Cg10005"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"
/translation="MTDPIEQAFERIRAEARRNGSVPLINKDAFRPPAPKGVK
RKGRASGLDGRQKRYVGAESLGSVTEIGRRGKDIAGMWTSSNEELYGAKIA
OHRVEMIKDKKLPICDSYAMATNLMMQROILQVIAKVGNNITELRIIGPQAS
WRKGPLVYKGRGPDITYG"
5354..7489
/gene="Cg10006"
5354..7489
CDS /gene="Cg10006"
/EC_number="5.99.1.3"
/note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:toprim domain
PF02518:Histidine kinase-, DNA gyrase B-, phycochrome-like
ATPase
TIGR01055:para: DNA topoisomerase IV, B subunit
TIGR01058:para: Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/transl_table=11
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="MRGTWGPKRVRMKRLYRISSECSLKVANTEHNYDASITILE
GLAVRRPQMYIGSTGRGLHLIWEVNVNSDVAAGATVETLLLEDGVOVD
DGRGIPDPMHSCAPTYQVYMTOLHAGKGSVAVSGAGHGVGISVYNALSTRREA
DIKLGKHWQNFESKYPDELISGNARGTGTTIRPPDAIEITTEPPEFISRLQ
EMALNKLGLITLTDNPADELELELAQGSTALELSDLEIDNETELVEERTDAK
KPKRERKKI.FHYENGLEDYVYHNRKTHPSIVSEKAGDHEVEVAMQNSVYK
ESVATPANTINTEGCTHEGFRSALISLNRVAREKLEKEANMLTGDCSEGSLA
VISVVDPOBEGQTKTKLQNTIEKISFVQMANEHIQHMENAPAAKVIINAVSA
QARLAARKADILVRKSGATDGLPGCLACRSKDPKSLYIVEGSGAKSKSRD
SMRQATLPLKGLINVERKALDKVLAETQATITLIGTSHDEPDKLRKHYIYM
ADADVDCGHATLTLTLTFRMPDLVAEGHYLAQPLAYTLKWRGEPGAYDEERD
EQULEDLAAGKINKDGIQRYKGLGEMASELNETTMDPTVAILKRVDTITDQRADE
LFSLMGDVVAARSFTTRNAKVRFLDI"
7830..8798
/gene="Cg10007"
7830..8798
CDS /gene="Cg10007"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="MIQYILIKIQSLKMAOREMLKTTLITDLGHVSTPTTAHP
LPDGSSTPVGIMASDNKDSQVLMVMPFGGYYRPLAALNKAGHVAIGELRGG

gene 3585..4769
/gene="Cg10004"
3585..4769
CDS /note="TIGR00611:recf: Recf protein"
/codon_start=1
/transl_table=11
/product="Recombinational DNA repair: ATPase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translation="WHIRSLLELDYRSWPELKVDLEPGITVFICNGNEKNTIVAGI
YLAHLSHRVSDAPLFAVRAENARVSAVAVQGRBLAHLIKPHAAQASLNTKV
TRRELIGVAVTLPFAPEDLALVGPSPARRRRLDIIATROPMAAGVADYKVLK
RNALLKTAATLALRRGYTEGGAALSLDITWDQGLARLGAEVMAAFALINELGRTY
EAYTTAPESRPAAVNYKTTIDQLSOFSEPDGITEATLTELAKROREIERGSS
VGRHDDVDIMLGDOPAKGPSHGETMSPLSIRIAEPNLKSGDIDPILIDDVSE
LDAGREKLVGIAQVEVQVLTAAVHDDPENLKVTLNQHTVTVQDITGRISILDV
Qp"
gene 4766..5302
/gene="Cg10005"
4766..5302
CDS /gene="Cg10005"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"
/translation="MTDPIEQAFERIRAEARRNGSVPLINKDAFRPPAPKGVK
RKGRASGLDGRQKRYVGAESLGSVTEIGRRGKDIAGMWTSSNEELYGAKIA
OHRVEMIKDKKLPICDSYAMATNLMMQROILQVIAKVGNNITELRIIGPQAS
WRKGPLVYKGRGPDITYG"
5354..7489
/gene="Cg10006"
5354..7489
CDS /gene="Cg10006"
/EC_number="5.99.1.3"
/note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:toprim domain
PF02518:Histidine kinase-, DNA gyrase B-, phycochrome-like
ATPase
TIGR01055:para: DNA topoisomerase IV, B subunit
TIGR01058:para: Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/transl_table=11
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="MRGTWGPKRVRMKRLYRISSECSLKVANTEHNYDASITILE
GLAVRRPQMYIGSTGRGLHLIWEVNVNSDVAAGATVETLLLEDGVOVD
DGRGIPDPMHSCAPTYQVYMTOLHAGKGSVAVSGAGHGVGISVYNALSTRREA
DIKLGKHWQNFESKYPDELISGNARGTGTTIRPPDAIEITTEPPEFISRLQ
EMALNKLGLITLTDNPADELELELAQGSTALELSDLEIDNETELVEERTDAK
KPKRERKKI.FHYENGLEDYVYHNRKTHPSIVSEKAGDHEVEVAMQNSVYK
ESVATPANTINTEGCTHEGFRSALISLNRVAREKLEKEANMLTGDCSEGSLA
VISVVDPOBEGQTKTKLQNTIEKISFVQMANEHIQHMENAPAAKVIINAVSA
QARLAARKADILVRKSGATDGLPGCLACRSKDPKSLYIVEGSGAKSKSRD
SMRQATLPLKGLINVERKALDKVLAETQATITLIGTSHDEPDKLRKHYIYM
ADADVDCGHATLTLTLTFRMPDLVAEGHYLAQPLAYTLKWRGEPGAYDEERD
EQULEDLAAGKINKDGIQRYKGLGEMASELNETTMDPTVAILKRVDTITDQRADE
LFSLMGDVVAARSFTTRNAKVRFLDI"
7830..8798
/gene="Cg10007"
7830..8798
CDS /gene="Cg10007"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="MIQYILIKIQSLKMAOREMLKTTLITDLGHVSTPTTAHP
LPDGSSTPVGIMASDNKDSQVLMVMPFGGYYRPLAALNKAGHVAIGELRGG

Query Match 100.0%; Score 324; DB 1; Length 340000;
Best Local Similarity 100.0%; Pred. No. 1.1e-88;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCCTGATTTCTCCTTTGTCGAGATGAGCATTAATTTT 60
277581 ATGACAACTGATTTCTCCTGATTTCTCCTTTGTCGAGATGAGCATTAATTTT 277640

61 GCGCTCCGGCGGCTTCCGTTCTTAATCCTTAAGCCCTAGCGATCAAAATTTGGGC 120
277641 GCGCTCCGGCGGCTTCCGTTCTTAATCCTTAAGCCCTAGCGATCAAAATTTGGGC 277700

121 AAAAAGCGATGTGATGTCGACGAGAAATCCTTGCAATTTGACCGATCAACGTTGCG 180
277701 AAAAAGCGATGTGATGTCGACGAGAAATCCTTGCAATTTGACCGATCAACGTTGCG 277760

181 ACCAATGCGATGATCTGAAGACTTAACCTTTGCTCATTTGCGGTTGCGATTACAGTG 240
277761 ACCAATGCGATGATCTGAAGACTTAACCTTTGCTCATTTGCGGTTGCGATTACAGTG 277820

241 GTGGCGCATCTTCTTGGCGGTCGACGACCTTTGTTAGACGTTGGCGCTGGCAACATCTTT 300
277821 GTGGCGCATCTTCTTGGCGGTCGACGACCTTTGTTAGACGTTGGCGCTGGCAACATCTTT 277880

301 TTTGTGACGATGATCTTTTTC 324
277881 TTTGTGACGATGATCTTTTTC 277904

RESULT 14
AX120085 349980 bp DNA linear PAT 11-MAY-2001
LOCUS

/gene="brnF"
/note="LIV-E transporter component 2; branched-chain amino acid export protein"
/codon_start=1
/transl_table=11
/product="brnF"
/protein_id="AA04686.1"
/db_xref="GI:21311381"
/translation="MKTQELHSLSVSPSKALEPDDKGRVYELAQELKSLAELGVPIGIAFGVLVQYGYEMWAPLPSGLIPAGSEMLVIALVGAAPLGIALTLVNVFHVAFSPFPLVQVWAPLAFVYFALIDEVAVYAPAPGMSAWRLISQIAHSYVWGGTGVIAELIPEIKGFALCSLFTVTLDSCKTKQIPSLLAGSFTI ALVIRPGALPAALIFLGILTRIFFGRAK"
1538..1993
/gene="ltp"
1538..1993
/gene="ltp"
/note="ltp"
/codon_start=1
/transl_table=11
/product="ltp-like regulator"
/protein_id="AA04687.1"
/db_xref="GI:21311382"
/translation="MKLSDIRALIAELSANARISNLADKVLTPGPCLRVORLEABGILGYSADIHPAVMNRGEVVDVTLNSPDSRPFSSVAQHDEVELHRLFGSPDVFVRIGVADLEAVEQFLSHIQTVPGIAKISRSFAMKVVPARPV"

ORIGIN

Query Match 100.0%; Score 324; DB 1; Length 2105;
Best Local Similarity 100.0%; Pred. No. 6,2e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACAGATTTCCCTGATTCCTGTTGTCGAGTATGAGTCACTACTTT 60
DB 711 ATGACACAGATTTCCCTGATTCCTGTTGTCGAGTATGAGTCACTACTTT 652
QY 61 GGGCTCCGGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGGGCG 120
DB 651 GGGCTCCGGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCAATTTGGGCG 592
QY 121 AAAATGGCGATGATGTCGACAGAGAACTTGGCACTTTGACGCGATCAACGTTTGGC 180
DB 591 AAAATGGCGATGATGTCGACAGAGAACTTGGCACTTTGACGCGATCAACGTTTGGC 532
QY 181 AGCAATGCGATGATCTGAAGACTTAACCTTGGTCTCAATTCGCGATTACAGTG 240
DB 531 AGCAATGCGATGATCTGAAGACTTAACCTTGGTCTCAATTCGCGATTACAGTG 472
QY 241 GGGGCGCATCTTCTGGCGGTGAGAGCACTTGTGAGCGCTGGGCGTGGACCATCGTT 300
DB 471 GGGGCGCATCTTCTGGCGGTGAGAGCACTTGTGAGCGCTGGGCGTGGACCATCGTT 412
QY 301 TTTGTTGACGACTGGTGAATCTTTTC 324
DB 411 TTTGTTGACGACTGGTGAATCTTTTC 388

RESULT 13
AP005274 340000 bp DNA linear BCT 08-AUG-2002
LOCUS Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
DEFINITION
ACCESSION AP005274 BA000036
VERSION AP005274.1 GI:21322764
KEYWORDS
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
REFERENCE 1 Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
AUTHORS Nakagawa, S.
TITLES Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 340000)
AUTHORS Nakagawa, S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa HAKKO Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@kanagen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
COMMENT This sequence is conducted by collaboration of Kyowa HAKKO Kogyo Co. Ltd. And Kitasato University.
FEATURES
source
1.340000
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:195627"
1..1575
/gene="Cg10001"
1..1575
/gene="Cg10001"
/note="PF00308:Bacterial dnaA protein
TIGR00362:dnaA: chromosomal replication initiator protein dnaA"
/codon_start=1
/transl_table=11
/product="Atpase involved in DNA replication initiation"
/protein_id="BAB97394.1"
/db_xref="GI:21322765"
/translation="MSQSSSLLETWROVADLTLSQDASGFDLPPTORAYMLTKPIAIDGIALVSTPNAKAVTENDLGAALRYVLSIRGRSSSLAVSEPBETPEAQERKYPDPAVISNSKAPQYEVGSGEASTSGMERTHSAPAEHPAIALDEPELATPQRIPRETPANPNREVSINPKYSEFVIGPFRFAAAVAABSPAKFNPLFISGSGSGKTHLAAVNGYAELOPGILIKVYSSBEFTDYLINSDRQETKRYRNLILMWDIOFLAGKEQOEPEFHEFNALQADKOIILSSDRPKQLTLEDLRTPEGGITLIDQPDLETRIALIMKAQODGTHVREVELIASRFSIRLEBLAGIRVSASSLINQIPDEMAIVLRDLIPSEDEITAPVIMEVTAEPFISVDTLGA GKTRAVAHARQALMTYCRELTDMSLPKIGDFGKCHITVMAADRLROEMTKRDTYDEIQULQILKSRGN"
/gene="Cg10002"
complement (1594..1920)
/gene="Cg10002"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translation="MIVNFEVYLMIDFKYRILSPVCAVSPAYPOFLGSCAQGKFRGITFPVTKLTQLQVCNTAVAKVVDNFSVNAFISYQNPCTTPKLNKDPEDSSBAAELNPFV"
2292..3476
/gene="Cg10003"
2292..3476
/EC number="2.7.7.7"
/note="PF00712:DNA polymerase III beta subunit, N-terminal domain
PF02767:DNA polymerase III beta subunit, central domain
PF02768:DNA polymerase III beta subunit, C-terminal domain
TIGR00663:dnaN: DNA polymerase III, beta subunit"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translation="WESQVTSFRVAREDDVTAVAWYAKNLPKTPQVYRAMLITADD EGSALGIDYDVSTVRLSAEVSQPKRIAVAGLSITGSLNPKPVDKIDSKAEVTCGSGREPLILPDDVPMPLKLPAGGSINACLFEVYSQVSAAGKDSLSMELGVSMELVQIMLAATDRFLALRTFEMANDPEINAVKLLPARTLIDNARSLDGLNDSIDLAVGDOVGREGLFEGVHNDRETTRLDADFPYIALPLKEHTALASVIALPVDATIRVSVLERNAGVILVHRSBGQVLTAGTAGAERTLPCAFGKELTIAPNGY LKQGLSVPTSRVAFVGTSPSRALMPEBENPANNENQIFQTPDTYTYLLMPVRL PG"

Query Match 100.0%; Score 324; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 5,8e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGATATGACGATCACTTTT 60
 DB 853 ATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGATATGACGATCACTTTT 912
 QY 61 GGGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATCAAAATTTGCGGC 120
 DB 913 GGGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATCAAAATTTGCGGC 972
 QY 121 AAAATGGCATGTGATGTCAGACAGAAATCCTTGCAATTTTGACCGCATCAAGCTTTCGC 180
 DB 973 AAAATGGCATGTGATGTCAGACAGAAATCCTTGCAATTTTGACCGCATCAAGCTTTCGC 1032
 QY 181 AGCAATGCGATGATCTGAAGCTTAACCTTGTGTCATTTGCGGTTGCGATTACAGTG 240
 DB 1033 AGCAATGCGATGATCTGAAGCTTAACCTTGTGTCATTTGCGGTTGCGATTACAGTG 1092
 QY 241 GTGGCGCATCTTCTTGCGCGGTGACGACCTTGTGAGCGTTGGCGCTGGACCATCGTT 300
 DB 1093 GTGGCGCATCTTCTTGCGCGGTGACGACCTTGTGAGCGTTGGCGCTGGACCATCGTT 1152
 QY 301 TTTGTTGACCTGTGATCTTTTC 324
 DB 1153 TTTGTTGACCTGTGATCTTTTC 1176

RESULT 11
 BD014994 1271 bp DNA linear PAT 27-JUN-2002
 LOCUS BD014994
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.

ACCESSION BD014994.1 GI:22555801
 VERSION BD014994.1
 KEYWORDS JP 2001169788-A/4.
 SOURCE unclassified
 ORGANISM unclassified
 TITLE 1 (bases 1 to 1271)
 AUTHORS Kemerikunehito, N., Sahm, H., Eggerling, L. and Pfeifferle, W.
 JOURNAL Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 PATENT: JP 2001169788-A 4 26-JUN-2001;
 DEGUSSA HUBIS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 COMMENT Corynebacterium glutamicum ATCC13032
 PN JP 2001169788-A/4
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KEMERIKUNEHITO, HERMANN SAHM, LOTHAR EGGERLING, WALTER PI
 PFEIFFERLE

PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08, C12N15/09, C12R1/15, (C12N1/21, C12R1/15), (C12P13/06, C12R1/15),
 PC C12N15/00,
 PC (C12N15/00, C12R1/15)
 CC brnF
 CC brnE
 CC key
 FT gene
 FT gene
 FT gene
 FT gene
 Location/Qualifiers
 location/Qualifiers
 1..1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 5,8e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGATATGACGATCACTTTT 60
 DB 853 ATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGATATGACGATCACTTTT 912
 QY 61 GGGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATCAAAATTTGCGGC 120
 DB 913 GGGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGATCAAAATTTGCGGC 972
 QY 121 AAAATGGCATGTGATGTCAGACAGAAATCCTTGCAATTTTGACCGCATCAAGCTTTCGC 180
 DB 973 AAAATGGCATGTGATGTCAGACAGAAATCCTTGCAATTTTGACCGCATCAAGCTTTCGC 1032
 QY 181 AGCAATGCGATGATCTGAAGCTTAACCTTGTGTCATTTGCGGTTGCGATTACAGTG 240
 DB 1033 AGCAATGCGATGATCTGAAGCTTAACCTTGTGTCATTTGCGGTTGCGATTACAGTG 1092
 QY 241 GTGGCGCATCTTCTTGCGCGGTGACGACCTTGTGAGCGTTGGCGCTGGACCATCGTT 300
 DB 1093 GTGGCGCATCTTCTTGCGCGGTGACGACCTTGTGAGCGTTGGCGCTGGACCATCGTT 1152
 QY 301 TTTGTTGACCTGTGATCTTTTC 324
 DB 1153 TTTGTTGACCTGTGATCTTTTC 1176

RESULT 12
 AF454053/c 2105 bp DNA linear BCT 27-JUN-2002
 LOCUS AF454053/c
 DEFINITION Corynebacterium glutamicum BrnE (brnE), BrnF (brnF), and Lrp-1-like
 regulator (Lrp) genes, complete cds.

ACCESSION AF454053.1 GI:21311379
 VERSION AF454053.1
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Corynebacteriaceae; Actinomycetales;
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 TITLE 1 (bases 1 to 2105)
 AUTHORS Kemeriknecht, N., Sahm, H., Yen, M.R., Patex, M., Sater, J.M.H. Jr. and
 Eggerling, L.
 JOURNAL Export of L-isoleucine from Corynebacterium glutamicum: a
 two-gene-encoded member of a new translocator family
 J. Bacteriol. 184 (14), 3947-3956 (2002)
 MEDLINE 22077265
 PUBMED 12081967
 REFERENCE 2 (bases 1 to 2105)
 Kemeriknecht, N., Eggerling, L. and Sahm, H.
 Direct Submission
 TITLE Submitted (07-NOV-2001) IBT-1, Forschungszentrum, Leo-Brandt Str.,
 JOURNAL Uelrich 52425, Germany
 Location/Qualifiers
 1..2105
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 complement(385..711)
 /gene="brnE"
 complement(385..711)
 /gene="brnE"
 /note="LIV-E transporter component 1; branched-chain amino
 acid export protein"
 /codon_start=1
 /transl_table=1
 /product="brnE"
 /protein_id="AA046685.1"
 /db_xref="GI:21311380"
 /translation="MTTDFSCILLVAVCAVITPALDAVPLIIPKRESQFVGKAM
 MPAIGIILITASTFRSNAIDKLTITGLIAVAITVAHLGGRRILLVGAGTIVF
 GLVNLFP"

FEATURES
 source
 gene
 CDS
 gene
 CDS
 complement(708..1463)
 /gene="brnF"
 complement(708..1463)

ORIGIN	
Query Match	100.0%; Score 324; DB 6; Length 1271;
Best Local Similarity	100.0%; Pred. No. 5, 8e-89;
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGACAACGTATTTCTCCTGTATTTCTCCTGTGTGTGCGACGATAGTGCAGTCATTACTTTT 60
DB	853 ATGACAACGTATTTCTCCTGTATTTCTCCTGTGTGTGCGACGATAGTGCAGTCATTACTTTT 912
QY	61 GCGCTCCGGGGGGGTTCCGTTCTTATCTTAAAGCCCTACGATGATACAAATTGTGGGC 120
DB	913 GCGCTCCGGGGGGGTTCCGTTCTTATCTTAAAGCCCTACGATGATACAAATTGTGGGC 972
QY	121 AAAATGGCGATGTGGATGCCAGCAGAAATCCTTGCCATTTTGACCGCATCAAGTTTCG 180
DB	973 AAAATGGCGATGTGGATGCCAGCAGAAATCCTTGCCATTTTGACCGCATCAAGTTTCG 1032
QY	181 AGCAATGGGATGATCTGGAAGACCTTAAACCTTTGATGAGCTTGGCGATTGCAATTAAGTG 240
DB	1033 AGCAATGGGATGATCTGGAAGACCTTAAACCTTTGATGAGCTTGGCGATTGCAATTAAGTG 1092
QY	241 GTGGCGCATCTTCTTGGCGGTCGACGACCTTTGTTGAGCGTGGCGCTGACCATGTTT 300
DB	1093 GTGGCGCATCTTCTTGGCGGTCGACGACCTTTGTTGAGCGTGGCGCTGACCATGTTT 1152
QY	301 TTTGTTGACCTGGTGAATCTTTTC 324
DB	1153 TTTGTTGACCTGGTGAATCTTTTC 1176
RESULT 9	
LOCUS	AX137714 1271 bp DNA linear PAT 30-MAY-2001
DEFINITION	Sequence 6 from Patent EP1096010.
ACCESSION	AX137714
VERSION	AX137714.1 GI:14273893
KEYWORDS	
SOURCE	
ORGANISM	Corynebacterium glutamicum
REFERENCE	Corynebacterium glutamicum
AUTHORS	Corynebacterium glutamicum
TITLE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
JOURNAL	1 Kennernecht, N., Eggeling, L., Sahm, H. and Pfeffeferle, W. Nucleotide sequences coding for branched-chain amino acids export proteins, method for isolating them and their use Patent: EP 1096010-A 6 02-MAY-2001;
FEATURES	1. .1271
source	location/Qualifiers
	/organism="Corynebacterium glutamicum"
	/mol_type="unassigned DNA"
	/db_xref="taxon:1718"
	/note="ATCC13032"
	101. .856
	/note="unnamed protein product; brnf"
	/codon_start=1
	/transl_table=1
	/protein_id="CAC39927.1"
	/db_xref="GI:14273894"
	/db_xref="RENTREMBL:CAC39927"
	/translation="MKTDEIHSSLEVEPSKAALEPPDKGRRRYEINAGLTSLAAGI
	VNPRFGIARGLVITQGYEMWAPLPSGLIEFGSTEMVIALVVGARLGAIALITLLI
	GVNRFIVAFSPFLHVNKNIAAPFYSPVFLIDEAVAVTPAPGASARLISMOAFR
	SLVVEGGLTGVAIAELIPEIIGLEFLCSPLVTLLDSCKTKOIPSLLAGSFTSL
	ALVPIGQALFAALILFLGLITIRYFELSKAK"
	853. .1179
	/note="unnamed protein product; brnf"
	/codon_start=1
	/transl_table=1
	/protein_id="CAC39928.1"
	/db_xref="GI:14273895"
	/db_xref="RENTREMBL:CAC39928"

Query Match	100.0%;	Score 324;	DB 6;	Length 1271;
Best Local Similarity	100.0%;	Pred. No. 5,8e-89;		
Matches 324;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
<p> /translation="MTTDSCLLVAVCAVITFPAVPEFLIKELRSSQFQKAA WNPAGLILATPSTRSNADLKTLLPGLIAVAITVAHLLGSRFTLLSVGAGTIVFV GLVNLPL" </p>				

RESULT 6
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1271)
AUTHORS Kemerhnecht, N., Sahm, H., Eggeling, L., and Pfeifferle, W.
TITLE Nucleotide sequences coding for the export of branched amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;
FEATURES
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCCGTAATCTTCCTGTTGTCGAGTATGACATCTACTTT 60
DB 853 ATGACACTGATTTCTCCGTAATCTTCCTGTTGTCGAGTATGACATCTACTTT 912
QY 61 GCGCTCCGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCACAATTTGTGGC 120
DB 913 GCGCTCCGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCACAATTTGTGGC 972
QY 121 AAATGCGCATGTGATGTCGACAGAGAACTCTTGCCATTTGACGCAATCAAGTTTGGC 180
DB 973 AAATGCGCATGTGATGTCGACAGAGAACTCTTGCCATTTGACGCAATCAAGTTTGGC 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTACCTTTGCTCATTTGCCGTTGCGATTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTACCTTTGCTCATTTGCCGTTGCGATTACAGTG 1092
QY 241 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGACCAATCGTT 300
DB 1093 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGACCAATCGTT 1152
QY 301 TTGTGACTGCTGATCTTTTC 324
DB 1153 TTGTGACTGCTGATCTTTTC 1176

RESULT 7
LOCUS AR391956 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1271)
AUTHORS Kemerhnecht, N., Sahm, H., Eggeling, L., and Pfeifferle, W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
FEATURES
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;

Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCCGTAATCTTCCTGTTGTCGAGTATGACATCTACTTT 60
DB 853 ATGACACTGATTTCTCCGTAATCTTCCTGTTGTCGAGTATGACATCTACTTT 912
QY 61 GCGCTCCGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCACAATTTGTGGC 120
DB 913 GCGCTCCGGCGGTTCCGTTCTTAATCTTAAAGCCCTTACGTAATCACAATTTGTGGC 972
QY 121 AAATGCGCATGTGATGTCGACAGAGAACTCTTGCCATTTGACGCAATCAAGTTTGGC 180
DB 973 AAATGCGCATGTGATGTCGACAGAGAACTCTTGCCATTTGACGCAATCAAGTTTGGC 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTACCTTTGCTCATTTGCCGTTGCGATTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTACCTTTGCTCATTTGCCGTTGCGATTACAGTG 1092
QY 241 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGACCAATCGTT 300
DB 1093 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGACCAATCGTT 1152
QY 301 TTGTGACTGCTGATCTTTTC 324
DB 1153 TTGTGACTGCTGATCTTTTC 1176

RESULT 8
LOCUS AX137709 1271 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1096010.
ACCESSION AX137709
VERSION AX137709.1 GI:14273886
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Kemerhnecht, N., Eggeling, L., Sahm, H., and Pfeifferle, W.
AUTHORS Nucleotide sequences coding for branched-chain amino acids export
TITLE Proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 1 02-MAY-2001;
DEGUSA AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
Location/Qualifiers
1..1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
101..856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/db_xref="REMBL:CAC39923"
/translation="NOKTOEIHSLSEVSPSKALEPDDKGRREYIAOGLKTSIAGL
GMVPIGAFGLIVIOGYEWMAALFSGLIFAGSTEMLVIALVVGALGLALATTL
VNPFRVAFSPFLHVNKPIARPSVPAIDEDVAVTAARPAQMSAMRLISQIAFH
SYNFGGLTGAVALIELPFEIKGEFPAKLSFYTLITDSCRTKQIPSLLAGSFTI
ALVPIPGALFPAALIFGLITIRFFDRAK"
853..1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMBL:CAC39924"
/translation="MTIDFSCILIVANCAVITPALRAVPLIILKPLRESQFVKMAN
WMPAGILAILTASTFRSNAIDKTLIFGLIAVAITVAHLIGSRITLISVAGTIVFV
GLVNLF"

CDS
101..856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/db_xref="REMBL:CAC39923"
/translation="NOKTOEIHSLSEVSPSKALEPDDKGRREYIAOGLKTSIAGL
GMVPIGAFGLIVIOGYEWMAALFSGLIFAGSTEMLVIALVVGALGLALATTL
VNPFRVAFSPFLHVNKPIARPSVPAIDEDVAVTAARPAQMSAMRLISQIAFH
SYNFGGLTGAVALIELPFEIKGEFPAKLSFYTLITDSCRTKQIPSLLAGSFTI
ALVPIPGALFPAALIFGLITIRFFDRAK"
853..1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMBL:CAC39924"
/translation="MTIDFSCILIVANCAVITPALRAVPLIILKPLRESQFVKMAN
WMPAGILAILTASTFRSNAIDKTLIFGLIAVAITVAHLIGSRITLISVAGTIVFV
GLVNLF"

QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

RESULT 4
 LOCUS BD014993
 DEFINITION BD014993 324 bp DNA linear PAT 27-AUG-2002
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014993
 VERSION BD014993.1 GI:22555800
 KEYWORDS JP 2001169788-A/3.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.

REFERENCE
 AUTHORS Kemerukunehito, N., Sahn, H., Eggering, L. and Pfeifferle, W.
 TITLE Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 JOURNAL Patent: JP 2001169788-A 3 26-JUN-2001;
 DEGUSA HUELS AG, FORSCHUNGSZENTRUM JOEHLICH GMBH
 COMMENT OS Corynebacterium glutamicum ATCC14752
 PN JP 2001169788-A/3
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KEMERUKUNEHITO, HERMANN SAHN, LOTHAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08//
 PC (C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
 PC C12N15/00,
 PC (C12N15/00, C12R1:15)
 CC bmrE
 CH Key
 FH key
 FT CDS Location/Qualifiers
 1..324 Location/Qualifiers
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 Location/Qualifiers
 1..324
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. Se-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCCGTATTCCTGTTGTCGAGTATGTCAGTCACTTACTTT 60
 DB 1 ATGACAACTGATTTCTCCGTATTCCTGTTGTCGAGTATGTCAGTCACTTACTTT 60
 QY 61 GCGCTCGGGCGGCTTCGTTCTTAATCCTTAAGCCCTTAAGTGAATCAAAATTGTGGC 120
 DB 61 GCGCTCGGGCGGCTTCGTTCTTAATCCTTAAGCCCTTAAGTGAATCAAAATTGTGGC 120
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 QY 181 AGCAATGCGATGATCTGAAGACTTAACTTTGGTCTCAATGCCGTGGATTAAGTG 240
 DB 181 AGCAATGCGATGATCTGAAGACTTAACTTTGGTCTCAATGCCGTGGATTAAGTG 240
 QY 241 GTGGCGCATCTTTCTGGCGGTGACAGCACTTGTGAGCGTTGGCGTGCACATCGTT 300
 DB 241 GTGGCGCATCTTTCTGGCGGTGACAGCACTTGTGAGCGTTGGCGTGCACATCGTT 300
 QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

RESULT 5
 LOCUS BD162491
 DEFINITION BD162491 324 bp DNA linear PAT 17-JAN-2003
 Novel polynucleotide.
 ACCESSION BD162491
 VERSION BD162491.1 GI:27868249
 KEYWORDS JP 2002191370-A/290.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.

REFERENCE
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 290 09-JUL-2002;
 KYOMA HAKKO KOGYO CO LTD
 COMMENT OS Corynebacterium glutamicum
 PN JP 2002191370-A/290
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 PI KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
 PI OZAKI
 PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
 C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
 PC 04, C12P13/08,
 PC C12P13/00, C12P13/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
 G01N33/56,
 PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15),
 PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
 PC C12N15/00,
 PC C12N15/00,
 CC Novel polynucleotide
 CH Key
 FH key
 FT source Location/Qualifiers
 1..324 Location/Qualifiers
 1..324 /organism="Corynebacterium glutamicum".
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source
 Location/Qualifiers
 1..324
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. Se-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCCGTATTCCTGTTGTCGAGTATGTCAGTCACTTACTTT 60
 DB 1 ATGACAACTGATTTCTCCGTATTCCTGTTGTCGAGTATGTCAGTCACTTACTTT 60
 QY 61 GCGCTCGGGCGGCTTCGTTCTTAATCCTTAAGCCCTTAAGTGAATCAAAATTGTGGC 120
 DB 61 GCGCTCGGGCGGCTTCGTTCTTAATCCTTAAGCCCTTAAGTGAATCAAAATTGTGGC 120
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTTGGCAATTTGACCGGATTAAGTTTGGC 180
 QY 181 AGCAATGCGATGATCTGAAGACTTAACTTTGGTCTCAATGCCGTGGATTAAGTG 240
 DB 181 AGCAATGCGATGATCTGAAGACTTAACTTTGGTCTCAATGCCGTGGATTAAGTG 240
 QY 241 GTGGCGCATCTTTCTGGCGGTGACAGCACTTGTGAGCGTTGGCGTGCACATCGTT 300
 DB 241 GTGGCGCATCTTTCTGGCGGTGACAGCACTTGTGAGCGTTGGCGTGCACATCGTT 300
 QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

FEATURES
source
Location/Qualifiers
1..324
/organism="unknown"
/mol_type="genomic DNA"

Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT
1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGAGATGTGAGTCATTACTTTT

61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120
61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180
121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240
181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300
241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324
301 TTGTGTGACTGTGAAATCTTTTC 324

RESULT 2
AX120374 324 bp DNA linear PAT 11-MAY-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX120374
Sequence 290 from Patent EP1108790.
AX120374
AX120374.1 GI:14037089

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE
JOURNAL
LOCATION/Qualifiers

1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Novel polymucleotides
Patent: EP 1108790-A 290 20-JUN-2001;
KIOWA HAKKO KOGYO CO., LTD. (JP)

1..324
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60
1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60

61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120
61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180
121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240
181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300
241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324
301 TTGTGTGACTGTGAAATCTTTTC 324

RESULT 3
AX137712 324 bp DNA linear PAT 30-MAY-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX137712
Sequence 4 from Patent EP1096010.
AX137712
AX137712.1 GI:14273891

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE
JOURNAL
LOCATION/Qualifiers

1 Kemmernecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
Protein, method for isolating them and their use
Patent: EP 1096010-A 4 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

1..324
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60
1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60

61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120
61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180
121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240
181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300
241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324
301 TTGTGTGACTGTGAAATCTTTTC 324

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324

301 TTGTGTGACTGTGAAATCTTTTC 324

RESULT 3
AX137712 324 bp DNA linear PAT 30-MAY-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX137712
Sequence 4 from Patent EP1096010.
AX137712
AX137712.1 GI:14273891

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE
JOURNAL
LOCATION/Qualifiers

1 Kemmernecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
Protein, method for isolating them and their use
Patent: EP 1096010-A 4 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

1..324
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60
1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60

61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120
61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180
121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240
181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300
241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324
301 TTGTGTGACTGTGAAATCTTTTC 324

RESULT 3
AX137712 324 bp DNA linear PAT 30-MAY-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX137712
Sequence 4 from Patent EP1096010.
AX137712
AX137712.1 GI:14273891

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE
JOURNAL
LOCATION/Qualifiers

1 Kemmernecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
Protein, method for isolating them and their use
Patent: EP 1096010-A 4 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

1..324
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60
1 ATGACAACTGATTTCTCTGATATTCCTGTTGTGACAGATGTGAGTCATTACTTTT 60

61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120
61 GCGCTCCGGCGGCTTCCTGTTTAATCCCTTAAGCCCTACGTAATCAATTGTGGGC 120

121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180
121 AAATGGCGATGTGATGACAGAGAAATCTTGCCATTTTGACCGCATCAACGTTTGGC 180

181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240
181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTCTCATTTGCCGTTACAGTG 240

241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300
241 GTGGCGCATCTTTCTGGCGGTGACGACCTTTGTGAGCGTTGGCGCTGGACCAATCGTT 300

301 TTGTGTGACTGTGAAATCTTTTC 324
301 TTGTGTGACTGTGAAATCTTTTC 324

RESULT 3
AX137712 324 bp DNA linear PAT 30-MAY-2001

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX137712
Sequence 4 from Patent EP1096010.
AX137712
AX137712.1 GI:14273891

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS
TITLE
JOURNAL
LOCATION/Qualifiers

1 Kemmernecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
Protein, method for isolating them and their use
Patent: EP 1096010-A 4 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 1373.24 Seconds
(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-4

Perfect score: 324
1 atgaactgattctctctg.....ttgacctggtgacatttc 324

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	324	100.0	324	AR391955	AR391955 Sequence
2	324	100.0	324	AX120374	AX120374 Sequence
3	324	100.0	324	AX137712	AX137712 Sequence
4	324	100.0	324	BD014993	BD014993 Nucleotide
5	324	100.0	324	BD162491	BD162491 Novel pol
6	324	100.0	1271	AR391953	AR391953 Sequence
7	324	100.0	1271	AR391956	AR391956 Sequence
8	324	100.0	1271	AX137709	AX137709 Sequence
9	324	100.0	1271	AX137714	AX137714 Sequence
10	324	100.0	1271	BD014991	BD014991 Nucleotide
11	324	100.0	1271	BD014994	BD014994 Nucleotide
12	324	100.0	2105	AF454053	AF454053 Coryneb
13	324	100.0	340000	AP005274	AP005274 Coryneb
14	324	100.0	349980	AX120085	AX120085 Sequence
15	38.2	11.8	65865	AY211269	AY211269 Campylob
16	36.6	11.3	4534	CFU25133	CFU25133
17	36	11.1	141605	AC013732	AC013732 Homo sapi
18	36	11.1	148332	AC079586	AC079586 Homo sapi
19	35.8	11.0	2040	NTA289862	NTA289862 Nicotiana
20	35.6	11.0	137678	AP005183	AP005183 Oryza sat
21	35.6	11.0	173270	AC024886	AC024886 Homo sapi
22	35.6	11.0	328197	AC109596	AC109596 Oryza sat
23	35.4	10.9	34494	AC006633	AC006633 Caenorhab
24	35.2	10.9	34494	AX655393	AX655393 Sequence
25	34.8	10.7	2000	AC104481	AC104481 Sus scrofa
26	34.8	10.7	86524	AC104481	AC104481 Sus scrofa
27	34.8	10.7	270748	AC121736	AC121736 Rattus no
28	34.6	10.7	110000	AC095209	AC095209 Rattus no
29	34.6	10.7	170475	AC136263	AC136263 Rattus no
30	34.6	10.7	179141	AC108890	AC108890 Bos tauru
31	34.6	10.7	224691	AC137364	AC137364 Rattus no
32	34.6	10.7	270850	AC095807	AC095807 Rattus no
33	34.4	10.6	189805	AC004169	AC004169 Homo sapi
34	34.4	10.6	190000	AC004479	AC004479 Homo sapi
35	34.2	10.6	1686	NOC00479	NOC00479 Homo sapi
36	34.2	10.6	12883	AE009221	AE009221 Agrobacte
37	34.2	10.6	12909	AE008188	AE008188 Agrobacte
38	34.2	10.6	13243	AE004602	AE004602 Pseudomon
39	34	10.5	150662	AC138550	AC138550 Danto rer
40	34	10.5	167753	AC108752	AC108752 Homo sapi
41	34	10.5	168332	AC053483	AC053483 Homo sapi
42	34	10.5	169030	AC125625	AC125625 Homo sapi
43	34	10.5	253370	AL929533	AL929533 Zebrafish
44	33.8	10.4	117645	AC144644	AC144644 Medicago
45	33.8	10.4	210791	CNS01D17	AL133467 Human chr

ALIGNMENTS

RESULT 1
LOCUS AR391955
DEFINITION Sequence 4 from patent US 6613545.
ACCESSION AR391955
VERSION AR391955.1 GI:40115726
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 324)
AUTHORS Kennertnecht,N., Salm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 4 02-SEP-2003

```

Db      66 GGANLASILLTLLTLLTSQILVGLSMRPLRSQPLRMRLGLGLTLLDEFFALTQYDRQ 125
QY      147 WSAWRLISMQLFHSYVWPGGLTGVIAIAELIP-FEIKLEFALCSLFTLLTLDSCRTKQ 205
Db      126 FRRWYALGVLPFYIAMNFTLGLGQNLPHLDQFGLDPSIVATFVA----- 174
QY      206 IPSLLAGSLFTIALVVIPOALFAAL-----LIFGLLTIRYFPL 246
Db      175 ----LIAPLVRLATVVCVAVSLFPCSVLFSYVWMEFALVAVAGLLGMSAGFI 221

```

RESULT 15

```

QTVI98      PRELIMINARY;      PRT;      231 AA.
AC      QTVI98;
DT      01-OCT-2003 (TrEMBLrel. 25, Created)
DT      01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein.
GN      HH0710.
OS      Helicobacter hepaticus.
OC      Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC      Helicobacteraceae; Helicobacter.
OX      NCBI_taxid=32025;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 51449 / 3B1;
RX      MEDLINE=22709201; PubMed=12810954;
RA      Suerbaum S., Josephans C., Sterzenbach T., Drescher B., Brandt P.,
RA      Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA      Holland K., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA      Schauer D.B., Shen Z., Weber J., Weber W., Fox J.G.;
RT      "The complete genome sequence of the carcinogenic bacterium
RT      Helicobacter hepaticus.";
RL      Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR      EMBL: AB017146; AAP7307.1; ".
KW      Hypothetical protein, Complete proteome.
SQ      SEQUENCE 231 AA; 25674 MW; 10CDA631C71EF9BD CRC64;

```

Query Match 15.3%; Score 194; DB 16; Length 231;
 Best Local Similarity 26.3%; Pred. No. 7.2e-08;
 Matches 51; Conservative 44; Mismatches 85; Indels 14; Gaps 6;

```

QY      44 LGWPIGIAFGLVIOGY-EMWAAPLFGSLFAGSTEMVIALVVGAPLGAIALTLTL 102
Db      23 LGYIFWGAVFGLLOKTKGYGAIWA-IMAVAIYGGITGFVAGLIAGAGLMSFVLVSM 81
QY      103 VNFHVFYAFSPFLHVVKNPIARFYSVFALIDEAVAVTAAR-----PAGWSAMRLISMQ 156
Db      82 INARQIFYSISMERFMGMKKRYMYISLDELFLNLNKSPTGVDKQWPFDFIAALN 141
QY      157 IAFHSYVVFGLTVAAIAELIPFEIKLEFALCSLFTLLTLDSCRTK-KQIPSLLAGLS 215
Db      142 ---QSYVVIQCTTGAAALGGKLEFNSGLDFWMSAIFVVFTEQMRNKTMRIPALF--GIT 196
QY      216 FTIALVVIPOALF 229
Db      197 ISLICLYIFGAGQF 210

```

Search completed: April 16, 2004, 06:55:50
 Job time : 53.6407 secs

OC Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_Taxid=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:H1 / CPT073 / ATCC 700928;
 RX MEDLINE=22386234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 Raako D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 Mayhew G.F., Rose D.C., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL: AE016765; AAN81687.1.
 DR PIR: C85916; C85916.
 DR PIR: H91071; H91071.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004471; AzIC.
 DR Pfam: PF03591; AzIC; 1.
 DR Pfam: PF03591; AzIC; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 245 AA; 26093 MW; 20A8F2D91B64D83 CRC64;
 Query Match 15.6%; Score 198; DB 16; Length 245;
 Best Local Similarity 29.1%; Pred. No. 3.6e-08;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;
 QY 34 QGLKTSLAGGMYPIGAFGLVLYQGYEWAAPLFSGLIIFAGSTEMLVIALVGAAPL 93
 DB 17 EGKDSLFIVSYIPVAFAGLNATRLGFSPLSEVFSCTIYAGSQFVTMLAAGSSL 76
 QY 94 GAILTLTVNFRHFVAFSPPLVIV---KNPFAFYSFALIDEXYATAR---PA 145
 DB 77 VVALTVAMDVRAHLVSPSRRIQLQKSKTALM--AFGLDEVFAATKLVENNR 134
 QY 146 GMSAMRLISMQIAFHSYVWF---GGLTGVAILLELFEIKGLFEALCSLFTVLTDSGR 201
 DB 135 RMSENMWIGIAFSSWSVFGTVAGVSGSLGLGYFAVENALGFMPLPAMFSLASQ 194
 QY 202 TKKQPSLLAGSTTLV-VIRGQALFA---ALL-IFLGULT--IRFFVIG 247
 DB 195 RKQ-----SLCVTALVAGLVTLFSLPAILAGIVCGGLTALIQAFWG 240
 RESULT 13
 Q3KBA PRELIMINARY; PRT; 247 AA.
 AC Q3KBA: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE YgaZ protein (Fragment).
 GN YGAZ.
 OS Erwinia chrysanthemi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OC NCBI_Taxid=556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3937;
 RA Reverchon S.;
 RT "Characterization of the Erwinia chrysanthemi multidrug resistance
 pump EmrAB."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ410307; CAC44347.1.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0006822; C:intracellular; IEA.
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004471; AzIC.

DR InterPro: IPR000524; HTH_GntR.
 DR Pfam: PF03591; AzIC; 1.
 DR PROSITE: PS00043; HTH_GNTR_FAMILY, 1.
 FT NON TER 1
 SQ SEQUENCE 247 AA; 26103 MW; 8958101295841873 CRC64;
 Query Match 15.5%; Score 197; DB 2; Length 247;
 Best Local Similarity 28.4%; Pred. No. 4.4e-08;
 Matches 67; Conservative 42; Mismatches 107; Indels 20; Gaps 7;
 QY 13 VSPKALPEDDKGRYREIAGLTKTSLAGLGMPIGAFGLVLYQGYEWAAPLFSG 72
 DB 3 VSSSQA---DARTTADASPAGSLFSLPIVIGYVAFAGLNATKGLFLEAIFLSC 58
 QY 73 LIPAGSTEMLVIALVGAAPLGAIALTLTVNFRHFVAFSPPLVVKR-PIAR-FYSVF 130
 DB 59 IIVAGSQFVTMLASGASIVWAALTVMAMDVRAHLVSPSRRIQKGLFLEAIFLSC 118
 QY 131 ALIDEAYATAPPA---GMSAMRLISMQIAFHSYVWFGLTGVAI---ALLPFIK 182
 DB 119 GLTDEVFAATKLVENNRMRSEPMWGLVALSWVAGTVI GAVFRNGPLEGYAVEA 178
 QY 183 GLSEALCSLFTVLTDSCTKQ-IPSLLAGLSTFIALVPIGQALFALIFVG 237
 DB 179 ALAFMLPALFLSLASFRRQSLVVAALGACUGLVSSIP-----AALLIGIG 229
 RESULT 14
 Q88KA5 PRELIMINARY; PRT; 230 AA.
 AC Q88KA5: 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Branched-chain amino acid transport protein AzIC.
 GN AzIC OR PP2385.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_Taxid=16048;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.B., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., Deboy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouli H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
 RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Medler H.,
 RA Lauber J., Stjepandic D., Hohnsbeil J., Straetz M., Helm S.,
 RA Kiewitz C., Eissen U., Timmis K.N., Duesterhoeft A., Tuelmler B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 metabolically versatile Pseudomonas putida KT2440."
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL: AE016783; AAN67998.1.
 DR TIGR: PP2385.
 DR GO: GO:0016021; C:Integral to membrane; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR004471; AzIC.
 DR Pfam: PF03591; AzIC; 1.
 KM Complete proteome.
 SQ SEQUENCE 230 AA; 24752 MW; 32C878FDC403A81 CRC64;
 Query Match 15.3%; Score 194; DB 16; Length 230;
 Best Local Similarity 25.5%; Pred. No. 7.1e-08;
 Matches 59; Conservative 43; Mismatches 101; Indels 29; Gaps 5;
 QY 29 RYIAQGLKTSLAGGMYPIGAFGLVLYQGYEWAAPLFSGLIIFAGSTEMLVIALV 88
 DB 6 RQAFHDAIAILPLSLAVFAGSLASMAIEANLSMOCQGLAIVFAAALIVAGIMK 65
 QY 89 GAFLGAIALTLTVNFRHFVAFSPPLVVKRPIA-RYSVFALIDEXYATAPR-AG 146

QY 150 WRISMOIAHSHYVWFGGLTGVAAIAELIP-FEIKGHEFALCSFTLTLDSCRTKKQIPS 208
 DB 136 WYALGAGLSFYTLFWNLATTLTGIVAGSLIPELNEMGLEFAVAATFAIVPTIKNAVLSS 195
 QY 209 ILLAGSFTIALV--IPGALFPAALLIFL 236
 DB 196 LVVA-LVGSVALFYQVBSGLMSITLAWL 224

RESULT 10

Q8DV59 PRELIMINARY; PRT; 232 AA.
 AC Q8DV59;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Putative integral membrane protein, possible branched-chain amino acid
 DE permease.
 DE SMU.388.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=2225063; PubMed=12397186;
 RA Ajdic D., Moshan W.M., McClughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najaf F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AEO14886; AANS8144.1; -
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0006810; P:Transport; IEA.
 DR InterPro; IPRO04471; AzIC.
 DR InterPro; IPRO0437; Prok_Lipoprot_S.
 DR Pfam; PF03591; AzIC; 1.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 232 AA; 24937 MW; 54654370CC0B47F6 CRC64;

Query Match 15.8%; Score 200.5; DB 16; Length 232;
 Best Local Similarity 28.7%; Pred. No. 2.2e-08;
 Matches 70; Conservative 44; Mismatches 91; Indels 39; Gaps 10;

QY 24 DKGYRREINQKSLTSLAAGLWYPIGIAFGLLVIOGYEWMAAPLFGSLIRAGSTEMLV 83
 DB 3 EKGEK---EGVAKSLPTLGLGYVSGLAGGVGANSGLTPFOMGLMSLTVAGSAQFWM 57
 QY 84 IALVGAAPLGAIALTLTVNFRHVAFVSFPLAV---VKNPVAFVSFVALI-DEAVA 138
 DB 58 CAMFVAGADPSIYMTVFILNLRN---FLMSLHATTFIKSLMQTICIGTLIDESYG 113
 QY 139 VTAR-----PAGSAMLLISMOIAHSHYVWFGGLTGVAAIAELIPF-EIKGHEFALCS 190
 DB 114 VLNEHYHAKOISTAMHGNNTGYLA---WLFNVLGTALGSVNPETIGLDPALIA 169
 QY 191 LFTVTL-----TLDSCTKKQIPSLLAGSFTIALVVIPOGALFPAALLIFL-----L 238
 DB 170 MFVSIFFSQALAMQAFKXKIGILLAVLSYFL-LVIVISLSLAVLSTLIGCFAGVL 228

QY 239 LTR 242
 DB 229 LDVA 232
 RESULT 11
 Q8X308 PRELIMINARY; PRT; 245 AA.
 ID Q8X308;
 AC Q8X308;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Hypothetical protein yga2.
 GN YGA2 OR C3235.

DT 01-MAR-2002 (TEMBLrel. 20, Created)
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Ort, hypothetical protein.
 GN 23983 OR EC3544.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_Taxid=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Poranousis K.,
 RA Apodaca U., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tode T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kubota S., Shiba T., Hattori M., Shingawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL; AEO05497; AAG57791.1; -
 DR EMBL; AP002562; BMB36967.1; -
 DR PIR; C85916; C85916.
 DR PIR; H91071; H91071.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0006810; P:Transport; IEA.
 DR InterPro; IPRO04471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KM Complete proteome.
 SQ SEQUENCE 245 AA; 26093 MW; 20AAFD291E644DB3 CRC64;

Query Match 15.6%; Score 198; DB 16; Length 245;
 Best Local Similarity 29.1%; Pred. No. 3.6e-08;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;

QY 34 QGLKTSIAAGLWYPIGIAFGLLVIOGYEWMAAPLFGSLIRAGSTEMLVIALVGAAPL 93
 DB 17 ECKDLSLPYIYISTIPAFGLNATPLGSPLESVFFSCITTAGSOPVITMTLAAGSL 76
 QY 94 GAIALTLTVNFRHVAFVSFPLAV---KNPVARFVSFVALIDEAVAVTAAR---PA 145
 DB 77 WVAALTYMADVDVHYVYGPSLSRSRIIRLQKSKTALM-AFGLTDEVFALAAAKIVRNRR 134
 QY 146 GMSAMRLISMOIAHSHYVWF---GSLTGVAAIAELIPFKEIKLEPALSFTLTLDSCRT 201
 DB 135 RNSENMIGIAFSWSWVFGIVIGAFSSGGLQCPANEAALGMLPALFMSFTLASFQ 194
 QY 202 TKQIBSLLAGSFTIALV-VIPGALFA---ALL-IFGLLT-IRYFGL 247
 DB 195 RKQ-----SLCVTAALVGAALGVTLSIVAILAGVCCGLTALIOAFWQ 240

RESULT 12
 Q8FE06 PRELIMINARY; PRT; 245 AA.
 ID Q8FE06;
 AC Q8FE06;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
 DE Hypothetical protein yga2.
 GN YGA2 OR C3235.


```

ID Q8G3T5 PRELIMINARY; PRT; 338 AA.
AC Q8G3T5;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Narrowly conserved hypothetical protein.
GN BL1669.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_Taxid=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC2 2705;
RX MEDLINE=2294977; PubMed=12381787;
RA Schell M.A., Karmaliatzou M., Snell B., Vilanova D., Berger B.,
RA Peesl G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AE014601; MAN25456.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004471; AzIC.
DR Pfam; PF03591; AzIC; 1.
DR Hypothetical protein; Complete proteome.
KM SEQUENCE 338 AA; 37538 MW; 805D6F5BA0C08A5 CRC64;

Query Match 16.9%; Score 214; DB 16; Length 338;
Best Local Similarity 29.9%; Pred. No. 2.6e-09;
Matches 64; Conservative 37; Mismatches 85; Indels 28; Gaps 8;

QY 44 LGWYPIGIAFGLVIOGYEWMWAAPLFSGLIIFAGSTEMVIALVGA-APLGNALTTLL 102
DB 99 LGFPLDASGYILMGTFGFSFWPMCMASATIFAGSEFVYVLLAFAFPLAGF-LLALM 157
QY 103 VNFHVFYAFSPFLHVVKNPIAF-----YVFALIDEAYAV--TAAPAGMS-AMR 151
DB 158 VNARHLFYGLSM-----LGKFKGLGWRKPYLFGMCDETFAINSTAKIPAGIDRGWF 209
QY 152 LISQVIAFHSGYVWFGITGVAIAPLPEIKGLFALCSFVLTLDSCRTKQIPSL-L 210
DB 210 YFWVTLLNQIYVWGTATLGLIGALHPFTGIDGDFVLTALFVLFDQMDLGRERRLSA 269
QY 211 LAGLSFTIALVPIPG-----QALFALLIFLGL 238
DB 270 VIGVLTSLACLLIFGANDFIMPIMWILITLVAL 303

RESULT 8
Q8Y223 PRELIMINARY; PRT; 242 AA.
ID Q8Y223
AC Q8Y223;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Hypothetical transmembrane protein RSC0513.
GN RSC0513 OR RS04992.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_Taxid=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangelot S.,
RA Ariat M., Billault A., Broctier P., Camus J.C., Catolico L.,
RA Chaudin M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Sigler P., Trebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;

```

```

RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
RL Nature 415:497-502(2002).
DR EMBL; AF646059; CAD1404.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004471; AzIC.
DR Pfam; PF03591; AzIC; 1.
KM Hypothetical protein; Complete proteome.
SO SEQUENCE 242 AA; 25196 MW; F22503FA454C28B CRC64;

Query Match 16.5%; Score 209; DB 16; Length 242;
Best Local Similarity 28.2%; Pred. No. 4.7e-09;
Matches 57; Conservative 38; Mismatches 95; Indels 12; Gaps 5;

QY 28 RRYEIAQGLTSLAAGIWMYPIGIAFGLVIOGYEWMWAAPLFSGLIIFAGSTEMVIALV 87
DB 5 RAHEFRAGVLTLPMLVGVVPGFIVGLVLAAGMPAMILAVAMSTIVFGASQMIIVOLW 64
QY 88 VGAAPLCAILTLTVNFRVFAFSPFLHVVKNPIA-RYSVFALIDEAYA-----VTA 141
DB 65 AGGAPALMAATVSMVNLRALYSASTAPALVHLPRWKLIVYLTDELFAAMRRVYN 124
QY 142 ARPAGMSA---WRLISQVIAFHSGYVWFGITGVAIAPLPEIKGLFALCSFVLTLL 197
DB 125 AQPDAQENTRHMFPLDAGVALMTSMQASTIVGALGAKVP-ATWPLDFLPLTFIAIV 183
QY 198 DSCRTKQIPSLIAGLSFTIA 219
DB 184 PSLTRPQL-APALTGAAIAYA 204

RESULT 9
Q871S6 PRELIMINARY; PRT; 236 AA.
ID Q871S6
AC Q871S6;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE AzIC family protein.
GN VPA0530.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_Taxid=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIND 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagemori K.,
RA Iijima Y., Nishima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
DR EMBL; AP005085; BAC61873.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004471; AzIC.
DR Pfam; PF03591; AzIC; 1.
KM Complete proteome.
SO SEQUENCE 236 AA; 25483 MW; 2288BE9F04DB15CB CRC64;

Query Match 16.0%; Score 203.5; DB 16; Length 236;
Best Local Similarity 27.1%; Pred. No. 1.3e-08;
Matches 57; Conservative 44; Mismatches 86; Indels 23; Gaps 8;

QY 44 LGWYPIGIA---FGLVIOGYEWMWAAPL---FSGLIIFAGSTEMVIALVGAAPLGA 96
DB 21 LAMPSTIAVLWGLLAGSPALDGLHPLLEGQALAIIFAGAGQVAMGMKAGAGLTTM 80
QY 97 ALTLTVNFRVFAFSPFLHVVKNPIA-RYSVFALIDEAYAVTAAPAG-----MSA 149
DB 81 LITTFITSRHFLYVSRSKSIAPLPKWRSLIGFLTLDELPAI-----AGHSDEQENR 135

```

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.,
 RA "The genome of *Methanosaerina acetivorans* reveals extensive metabolic
 RT and physiological diversity." ;
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AB010501; AA06804.1; -
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome.
 SQ SEQUENCE 291 AA; 32067 MW; AFA5100B00B1B8 CRC64;
 Query Match 23.7%; Score 301; DB 17; Length 291;
 Best Local Similarity 32.6%; Pred. No. 2.4e-16;
 Matches 79; Conservative 38; Mismatches 97; Indels 28; Gaps 6;
 QY 24 DKGRRYRIAG-----LKTSLAAGLGMYPGIAFGILVIGYEWMAAPLPSGLIF 75
 DB 23 EKEVRRYTRGGDRPLFLSALKTTPVPLGYIPGLMAFGFLMDAGYHWIYAFMSLLIY 82
 QY 76 AGSTEMLVIALVGAAPGALATLTLVNFHVFYAFSPFLHVKNPPIARF-----Y 127
 DB 83 AGAGCPFLAVALLAAGAGLPEFVIALTLNLHRAFGLSL-----LDKSDIGKVKPY 134
 QY 128 SVEFALIDEAVY--TAAPAGMSAMRLISMQIAF-HSYWVFGTLGVAAIEIPFEIKGL 184
 DB 135 LIFALTDYVALLTHVEVPAGSKARFFYISALDHFWVAGSVLGAAGSLNLMNIEGM 194
 QY 185 EFALCSLPVTLTDS-CRTKQIPSLIAGSFTIALVVIQGFALLIFLGLITRY 243
 DB 195 AFVLTALFVLTIEQYFNSIRFPVAAVGAAGASLILSPENMLISILIGTLIARE 254
 QY 244 PF 245
 DB 255 KF 256
 RESULT 5
 Q88325 PRELIMINARY; PRT; 244 AA.
 AC Q88325;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Branched-chain amino acid transport protein.
 GN AZIC OR MM3362.
 OS Methanosaerina mazel (*Methanosaerina frisia*).
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosaerinales; Methanosaeriniaceae; Methanosaerina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Geel / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=2210827; PubMed=12125824;
 RA Depierremer U., Ucham A., Hartech T., Merkl R., Schmitz R.A.,
 RA Martineau A., Henne A., Wietzer A., Benner S., Jacob C.,
 RA Buesgen H., Lénard T., Christmann A., Boemcke M., Steckel S.,
 RA Bhatnagar A., Lykidis A., Overbeek R., Klein H.-P., Gunsalus R.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of *Methanosaerina mazel*: evidence for lateral gene
 RT transfer between Bacteria and Archaea." ;
 RT J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AB013585; AA032858.1; -
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome.
 SQ SEQUENCE 244 AA; 26655 MW; AC8A74DA9EBCC60 CRC64;

Query Match 22.5%; Score 285; DB 17; Length 244;
 Best Local Similarity 32.8%; Pred. No. 3.9e-15;
 Matches 76; Conservative 40; Mismatches 86; Indels 28; Gaps 9;
 QY 23 DKGRRYRIAG-----LKTSLAAGLGMYPGIAFGILVIGYEWMAAPLPSGLIF 81
 DB 12 DDK-----DVFGLAKTTVPVPLGYIPGLMAFGFLMDAGYHWIYAFMSLLVAGSQF 66
 QY 82 LVIALVGAAPGALATLTLVNFHVFYAFSPFLHVKNPPIARF-----YVFALI 133
 DB 67 LAVALLAAGAGLTERVIALTLNLHRAFGLSL-----LEKSDVGKVKPYLIFALT 118
 QY 134 DEAYAV--TAAPAGMSAMRLISMQIAF-HSYWVFGTLGVAAIEIPFEIKGFALCS 190
 DB 119 DETYALLTTTEVPKQGSRRFFYIADLHVLWITGSVLGALLSLDLNLEGMAFVLT 178
 QY 191 LFTVTLTDS-CRTKQIPSLIAGSFTIALVVIQGFALLIFLGLIT 241
 DB 179 LFTVLTIEQYFNSIRFPVAAVGAAG-TLSILSPENML-LISILIGTLIL 228
 RESULT 6
 Q9CNK7 PRELIMINARY; PRT; 240 AA.
 ID Q9CNK7;
 AC Q9CNK7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PM0422.
 GN PM0422.
 OS *Pasteurella multocida*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pm70;
 RX MEDLINE=2145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittem T.S., Kapur V.;
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70." ;
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL EMBL; AB006078; AA02506.1; -
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 240 AA; 26864 MW; 63A1770AC2198BBE CRC64;
 Query Match 18.7%; Score 237; DB 16; Length 240;
 Best Local Similarity 31.3%; Pred. No. 2.7e-11;
 Matches 68; Conservative 41; Mismatches 94; Indels 14; Gaps 8;
 QY 45 GMPYIGAFGLVIGYEWMAAPLPSGLIFAGSTEMLVIALVGA-APGALATLTLV 103
 DB 26 GFLPLGIAVGYMAGLGEALYPTLMLFYAGSVEITVAGLVAPSPLN-VLLITLMV 84
 QY 104 NFRHVFYAFS-----FPLHVNPNPIARFYSVFALIDEAYAVT--AARPAGMS-AMRLISMQ 156
 DB 85 SGRIQFQISMLKXGAVLGKK--RWYLTISTVDEAFSINVAKVVEGIERGVFFVS 141
 QY 157 IAHSHYVFGTLGVAAIEIPFEIKGFALCSFTVTLTDS-CRTKQIPSLIAGLS 215
 DB 142 FYLOFYWVIGAIGALFSGSLIFPLSGIEFMTALFLVIFAEQCKERSHESALIGLGA 201
 QY 216 FTIALVVIQGFALLIFLGLITIRYFPL-GPAK 251
 DB 202 FTIALVVKTYFLIPTLIGIFALTIRVKSATLAK 238
 RESULT 7
 Q8G3T5

Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVPSPKALPEPDKGYRREYINQGLKTSIAGLGMYPGIAFGLLIYQ 60
DB 1 MOKTQEHSSLEVPSPKALPEPDKGYRREYINQGLKTSIAGLGMYPGIAFGLLIYQ 60
QY 61 GYEMWAPLPSGLIFASSTEMLVAVVGAAPGALATLTLVNFHVFYAFSPPLHYK 120
DB 61 GYEMWAPLPSGLIFASSTEMLVAVVGAAPGALATLTLVNFHVFYAFSPPLHYK 120
QY 121 NPIRFFSVFALIDEAVAVTAARAGMSANRLISMQAFHSYVWFGSLGVAAELIPE 180
DB 121 NPIRFFSVFALIDEAVAVTAARAGMSANRLISMQAFHSYVWFGSLGVAAELIPE 180
QY 181 IKGLEFALCSLFTVLTLDSCRTKQIPSLLAGISFTIALVIFGOALFALLIFGLLT 240
DB 181 IKGLEFALCSLFTVLTLDSCRTKQIPSLLAGISFTIALVIFGOALFALLIFGLLT 240
QY 241 IRYFFLGKAK 251
DB 241 IRYFFLGKAK 251

RESULT 2

Q8PSW8 PRELIMINARY; PRT; 238 AA.

ID Q8PSW8
AC Q8PSW8
RT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN CE2926.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacterium.
NCBI_TaxID=152794;

QY 11 SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005224; BAC19736.1; -
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR DR GO:0006810; P:Transport; IEA.
DR InterPro; IPR004471; AzIC.
DR Pfam; PF03591; AzIC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 238 AA; 25739 MW; 02BDEDBAB33A31 CRC64;

Query Match 35.3%; Score 448.5; DB 16; Length 238;
Best Local Similarity 42.3%; Pred. No. 3e-28;
Matches 91; Conservative 42; Mismatches 79; Indels 3; Gaps 3;

QY 31 EIAQGLKTSIAGLGMYPGIAFGLLIYQGYEMWAPLPSGLIFASSTEMLVAVV-G 89
DB 31 EIAQGLKTSIAGLGMYPGIAFGLLIYQGYEMWAPLPSGLIFASSTEMLVAVV-G 89
QY 8 EIGGIRFSAVGLVPLGLAVGSLVMMWPFISITYAGSMYLAIGLAA 67
DB 8 EIGGIRFSAVGLVPLGLAVGSLVMMWPFISITYAGSMYLAIGLAA 67
QY 90 AAPGALATLTLVNFHVFYAFSPPLHYKXNPPIAFYSFALIDEAVATAAPG-WS 148
DB 90 AAPGALATLTLVNFHVFYAFSPPLHYKXNPPIAFYSFALIDEAVATAAPG-WS 148
QY 68 VGPSFAL-VTGFVNFRHIFGLTFPRNAIRSKGRAYSTVALTDNAVIAASRPGEIS 126
DB 68 VGPSFAL-VTGFVNFRHIFGLTFPRNAIRSKGRAYSTVALTDNAVIAASRPGEIS 126
QY 149 AMRLISMQAFHSYVWFGSLGVAAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
DB 149 AMRLISMQAFHSYVWFGSLGVAAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
QY 127 GTRLLTLQIVCQSLWFGPIGAVAGALPDGKMGMEFALTALFVLAWEAFNNNDWSL 186
DB 127 GTRLLTLQIVCQSLWFGPIGAVAGALPDGKMGMEFALTALFVLAWEAFNNNDWSL 186
QY 209 LLLAGLSFTIALVIFGOALFALLIFGLITRYF 243
DB 209 LLLAGLSFTIALVIFGOALFALLIFGLITRYF 243
QY 187 PLFAVAVLAVSGFAPEQMLVLTLYFILLRYF 221
DB 187 PLFAVAVLAVSGFAPEQMLVLTLYFILLRYF 221

RESULT 3

Q8NL66 PRELIMINARY; PRT; 237 AA.

ID Q8NL66
AC Q8NL66
RT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Predicted branched-chain amino acid permease (azalucine resistance).
GN CG13083.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacterium.
NCBI_TaxID=1718;

QY 11 SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005283; BAC00477.1; -
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR DR GO:0006810; P:Transport; IEA.
DR InterPro; IPR004471; AzIC.
DR Pfam; PF03591; AzIC; 1.
KW Complete proteome.
SQ SEQUENCE 237 AA; 25867 MW; 193EB1182E0CF923 CRC64;

Query Match 33.8%; Score 428.5; DB 16; Length 237;
Best Local Similarity 40.1%; Pred. No. 1.2e-26;
Matches 87; Conservative 46; Mismatches 81; Indels 3; Gaps 3;

QY 31 EIAQGLKTSIAGLGMYPGIAFGLLIYQGYEMWAPLPSGLIFASSTEMLVAVV-G 89
DB 31 EIAQGLKTSIAGLGMYPGIAFGLLIYQGYEMWAPLPSGLIFASSTEMLVAVV-G 89
QY 8 EIRGGISITLVGGLIPGLAGLMDVGTGAMWMTPIFSVITYAGSMYLAIGMTAG 67
DB 8 EIRGGISITLVGGLIPGLAGLMDVGTGAMWMTPIFSVITYAGSMYLAIGMTAG 67
QY 90 AAPGALATLTLVNFHVFYAFSPPLHYKXNPPIAFYSFALIDEAVATAAPG-WS 148
DB 90 AAPGALATLTLVNFHVFYAFSPPLHYKXNPPIAFYSFALIDEAVATAAPG-WS 148
QY 68 IGPSA-AVAGFVNFRHIFGLTFPRRIKSGAGRAYSTVALTDESVAIVASRPDIS 126
DB 68 IGPSA-AVAGFVNFRHIFGLTFPRRIKSGAGRAYSTVALTDESVAIVASRPDIS 126
QY 149 AMRLISMQAFHSYVWFGSLGVAAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
DB 149 AMRLISMQAFHSYVWFGSLGVAAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
QY 127 GTRLLTLQIVCQSLWFGPIGAVAGALPDGKMGMEFALTALFVLAWEAFNNNDWSL 186
DB 127 GTRLLTLQIVCQSLWFGPIGAVAGALPDGKMGMEFALTALFVLAWEAFNNNDWSL 186
QY 209 LLLAGLSFTIALVIFGOALFALLIFGLITRYF 245
DB 209 LLLAGLSFTIALVIFGOALFALLIFGLITRYF 245
QY 187 PLFAVAVLAVSGFAPEQMLVLTLYFILLRYF 223
DB 187 PLFAVAVLAVSGFAPEQMLVLTLYFILLRYF 223

RESULT 4

Q8TKHO PRELIMINARY; PRT; 291 AA.

ID Q8TKHO
AC Q8TKHO
RT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Branched chain amino acid transport protein AzIC.
GN MAJ437.
OS Methanosaetia acetivorans.

QY 11 SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
R Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
Allen N., Naylor J., Stange-thomann N., Deatellano K., Johnson R.,
Linton L., McEwan P., McKernan K., Talamas D.E., Graham D.A., Ye W.,
Zimmer A., Barber R.D., Camm I., Graham D.E., Graham D.A., Guss A.M.,
Hedderich R., Ingram-Smith C., Kuetner H.C., Kzyzski J.A.,
Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 49.6407 Seconds

(without alignments)
1595.367 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269

Sequence: 1 MORTGHEHSLSEVPSKAL.....LLIFGLTINRYFKNAK 251

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	16	Q8NTP8
2	448.5	35.3	251	16	Q8NTP8
3	428.5	33.8	237	16	Q8NTP8
4	301	23.7	231	17	Q8NTP8
5	285	22.5	244	17	Q8NTP8
6	237	18.7	240	16	Q8NTP8
7	214	16.9	238	16	Q8NTP8
8	209	16.5	242	16	Q8NTP8
9	203.5	16.0	236	16	Q8NTP8
10	200.5	15.8	232	16	Q8NTP8
11	198	15.6	245	16	Q8NTP8
12	198	15.6	245	16	Q8NTP8
13	197	15.5	247	2	Q93KB4
14	194	15.3	230	16	Q8KAS
15	194	15.3	231	16	Q7V198
16	194	15.3	245	16	Q83JZ6

17	192	15.1	250	16	Q8RE08
18	186	14.7	237	16	Q8DE60
19	185	14.6	230	16	Q8E349
20	184.5	14.5	237	16	Q7M08
21	184.5	14.5	237	16	Q7M08
22	182	14.3	241	16	Q8ISD8
23	181.5	14.3	241	16	Q8IFC0
24	181.5	14.3	256	16	Q82R05
25	181	14.3	234	16	Q8UCR6
26	181	14.3	239	16	Q8KXU8
27	177.5	14.0	218	16	Q87114
28	177	13.9	233	16	Q88Y19
29	176.5	13.9	235	16	Q83107
30	176.5	13.9	242	16	Q88K06
31	176	13.8	261	16	Q88T22
32	175	13.8	252	16	Q81276
33	175	13.8	257	16	Q82BX0
34	174.5	13.8	235	16	Q8CF68
35	174.5	13.8	237	16	Q8K802
36	169	13.3	235	16	Q8Y761
37	168	13.2	241	16	Q8CUI8
38	167	13.2	230	16	Q82V77
39	167	13.2	230	16	Q8DX08
40	166	13.1	235	16	Q82BR4
41	163.5	12.9	219	16	Q8XGA3
42	163	12.8	238	16	Q82N22
43	161.5	12.7	230	16	Q8C0K2
44	161	12.7	224	2	Q8YWS9
45	161	12.7	224	16	Q8FYU5

ALIGNMENTS

RESULT 1

Q8NTP8 PRELIMINARY; PRT; 251 AA.

AC Q8NTP8;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Predicted branched-chain amino acid permease (azalaucine resistance) (Brnf).
GN CG10258 OR BRNF.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxId=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2077265; PubMed=12081967;
RA Kempe-Knecht N., Sahn H., Yernool R., Patek M., Sailer M.H. Jr., Eggeling L.;
RT "Export of L-isoleucine from Corynebacterium glutamicum: a Two-Gene-Encoded Member of a New Translocase Family.";
RL J. Bacteriol. 184:3947-3956(2002).
DR EMBL: AP005274; BAB97651.1;
DR EMBL: AF454053; AA46686.1;
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro: IPR004471; AzIC.
DR Pfam: PF03591; AzIC; I.
KM Complete proteome.
SQ SEQUENCE 251 AA; 27333 MW; 898F20D94320D66 CRC64;
Query Match 100.0%; Score 1269; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 6.1e-94;

```

FT VARIANT 67 67 N -> S.
FT VARIANT 102 102 P -> S (IN C/SH2 ANTIGEN).
FT VARIANT 153 153 P -> T (IN RHE KH).
FT VARIANT 225 225 P -> A (IN E/RHS ANTIGEN).
FT VARIANT 232 232 P -> E (IN RHE FM).
FT VARIANT 237 237 P -> V (IN RHE FM).
FT VARIANT 244 244 P -> V (IN VS ANTIGEN).
FT VARIANT 244 244 P -> V (IN VS ANTIGEN).
FT CONFLICT 11 11 C -> L (IN REF. 7).
FT CONFLICT 52 52 D -> G (IN REF. 4).
FT CONFLICT 60 60 G -> C (IN REF. 4).
SQ SEQUENCE 416 AA; 45429 MM; DC7AF97DC0DCFF1F CRC64;

Query Match 7.6%; Score 96.5; DB 1; Length 416;
Best Local Similarity 21.7%; Pred. No. 1.3;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 AALEPDDGQ-VRRYRIAGLKTSLAAGMP-----IGAFGLVIOGYEMWAP 68
DB 35 ASLE-DQKGLVASVGVODLTMAALGLFLTSNRRSSMSVAFNLFMLALGVQW--AI 91
QY 69 IFSG-----LIFAGST-----EMVIALVGAAPLGA 95
DB 92 LLDGFLSOPPKGVITLFSIRLATMSAVSLISAGAVLKVLAQLVVMVLVEVTAAGT 151
QY 96 IAL-----TLLVNFRRVFAFS--PELHV--VKNPI-----ARFYSVPAFI 133
DB 152 LEMVISNINFTYHMLLRH-FYFPAVYFGLTVAKCLPRLPGTENDORATIPSLML 210
QY 134 DEAYVATARPAGMSAMRLISMQI-----AFHSYWGGLTGVAIAELIPEIKG 183
DB 211 GALTFL-----WFMFSPNSPLRSPIQKNAEMNTY-----ALAVSVTAISG 254
QY 184 LEPAICSLFVTLTDSCKTKQIPSLILA-GLSFTIALVVIPOALFALLIFGLTIT 241
DB 255 SSIAHQRKISMT-----YHSAVLAGVAVGTSCHLIPSPMLAVLGLVAGLISI 305

RESULT 15
CLC4 MOUSE STANDARD; PRT; 747 AA.
AC 061418;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chloride channel protein 4 (CLC-4).
OS CLC4 OR CLC4-2 OR CLC4.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=CS7BL/6J; TRISUB=Retina;
RX MEDLINE=95400329; PubMed=7670496;
RA Rugarli B.I., Adler D.A., Borsani G., Tsuchiya K., Franco B.,
RA Hauge X., Diatche C., Chapman V., Ballabio A.;
RT "Different chromosomal localization of the Clcn4 gene in Mus spretus
RT and C57BL/6J mice."
RI Nat. Genet. 10:466-471(1995).
CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC trans epithelial transport.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -1- SIMILARITY: Contains 2 CBS domains.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z49916; CAA90150.1; -.
DR PIR; I48294; I48294.
DR MGD; MGI:104571; Clcn4-2.
DR InterPro; IPR000644; CBS domain.
DR InterPro; IPR01807; CL-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
DR Ionic channel; Ion transport; Chloride channel; Chloride;
KW Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 454 474 POTENTIAL.
FT TRANSMEM 495 515 POTENTIAL.
FT TRANSMEM 522 542 POTENTIAL.
FT TRANSMEM 586 646 CBS 1.
FT DOMAIN 684 734 CBS 2.
SQ SEQUENCE 747 AA; 83974 MM; B5486AA1B0721144 CRC64;

Query Match 7.5%; Score 95; DB 1; Length 747;
Best Local Similarity 25.6%; Pred. No. 3;
Matches 40; Conservative 25; Mismatches 57; Indels 34; Gaps 8;

QY 80 EMVIALVVG-----AAPLGAIALTLLVNFRRVFAFSPLHVKNPIRFPYSVPAFI 134
DB 241 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--YFPLKTL-----WMSFFPALV 287
QY 135 EAYVATARPAGMSAMRLISMQIAFHSYWGGLTGVAIAELIPEIKG:EPALC-SLFV 193
DB 288 AAFTRISINPFGNS--RLVLFVEYHTPWY-----MABLFFILLSVFGGLNGTLFT 337
QY 194 TLTLDSCKTKK-----OIPSLLAGLSFTIALVVI 224
DB 338 RCNIAMCRKRKTTLLGRYPVLEVAIVATAIVAVP 373

```

Search completed: April 16, 2004, 06:53:44
Job time : 16.9833 secs

RN [2] SEQUENCE FROM N.A.
 RP MEDLINE=91058523; PubMed=2123099;
 RX Avent N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;
 RA "cDNA cloning of a 30 kDa erythrocyte membrane protein associated
 RT with Rh (Rhesus)-blood-group-antigen expression.";
 RL Biochem. J. 271:821-825(1990).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=93216283; PubMed=7916743;
 RA Kaji E., Umenishi F., Iwamoto S., Ikemoto S.;
 RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated
 RL with the Rh blood group system.";
 RL Hum. Genet. 91:157-162(1993).
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORMS RHIV, RHVI AND RHVII).
 RP TISSUE=Bone marrow;
 RC MEDLINE=92360855; PubMed=1379850;
 RX le van Kim C., Cherif-zahar B., Raynal V., Mouro I., Lopez M.,
 RA Cartron J.-P., Colin Y.;
 RA "Multiple Rh messenger RNA isoforms are produced by alternative
 RT splicing.";
 RT Blood 80:1074-1078(1992).
 RN [5]
 RN SEQUENCE FROM N.A., AND VARIANTS THR-153; GLU-232 AND VAL-237.
 RP Kaestliwase K., Ishikawa Y., Uchikawa M.;
 RA "The serological profile and molecular basis of the RHE variants
 RT (RHEFM and RHEKH) in Japanese.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE OF 1-32.
 RP MEDLINE=89131633; PubMed=3145980;
 RX Avent N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J.,
 RA Kumpel B.;
 RA "Protein-sequence studies on Rh-related polypeptides suggest the
 RT presence of at least two groups of proteins which associate in the
 RL human red-cell membrane.";
 RL Biochem. J. 256:1043-1046(1988).
 RN [7]
 RN SEQUENCE OF 1-16.
 RP MEDLINE=88294325; PubMed=3135863;
 RX Bloy C., Blanchard D., Dahr W., Beyreuther K., Salmon C.,
 RA Cartron J.-P.;
 RA "determination of the N-terminal sequence of human red cell Rh(D)
 RT polypeptide and demonstration that the Rh(D), (C), and (E) antigens
 RL are carried by distinct polypeptide chains.";
 RL Blood 72:661-666(1988).
 RN [8]
 RN SEQUENCE OF 1-10 FROM N.A.
 RP MEDLINE=94451582; PubMed=8188244;
 RX Cherif-zahar B., le van Kim C., Rouillac C., Raynal V., Cartron J.-P.,
 RA Colin Y.;
 RA "Organization of the gene (RHEB) encoding the human blood group
 RT RhCBE antigens and characterization of the promoter region.";
 RL Genomics 15:68-74(1994).
 RN [9]
 RN VARIANTS BLOOD GROUP C AND E.
 RP MEDLINE=94035121; PubMed=8220426;
 RX Mouro I., Colin Y., Cherif-zahar B., Cartron J.-P., le van Kim C.;
 RA "Molecular genetic basis of the human Rhesus blood group system.";
 RL Nat. Genet. 5:62-65(1993).
 CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
 CC HAVE A TRANSPORT OR CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: Integral membrane protein.
 CC Event=Alternative splicing; Named Isoforms=4;
 CC Name=RHI;
 CC IsoId=p18577-1; Sequence=Displayed;
 CC Name=RHIV;
 CC IsoId=p18577-2; Sequence=VSP_005703, VSP_005704;
 CC Name=RHVI;
 CC IsoId=p18577-3; Sequence=VSP_005702, VSP_005705;
 CC Name=RHVII;
 CC

CC		IsoId=I18577-4; Sequence=VSP_005701;	
CC	-1-	TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING ERYTHROID CHARACTERS.	
CC	-1-	POLYMORPHISM: RhCE and RHD are responsible for the RH blood group system. The molecular basis of the B-Rh3/=Rh5 blood group antigens is a single variation in position 225; Pro-225 corresponds to Rh3 and Ala-225 to Rh5. The molecular basis of the CeRh2/Ce-th4 blood group antigens is a single variation in position 102; Ser-102 corresponds to Rh2 and Pro-102 to Rh4.	
CC	-1-	SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHD.	
CC	-----		
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC between the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-eib.ch).	
CC	-----		
DR	EMBL;	M34015; AAA36567.1; -	
DR	EMBL;	X54534; CAA38401.1; -	
DR	EMBL;	S57967; AAB26080.1; -	
DR	EMBL;	X63095; CAA44809.1; -	
DR	EMBL;	X63096; CAA44810.1; -	
DR	EMBL;	X63098; CAA44812.1; -	
DR	EMBL;	AB018644; BA33927.1; -	
DR	EMBL;	AB018645; BA33928.1; -	
DR	PIR;	A30405; AAD14061.1; -	
DR	Gene;	HGNC:10008; RHCE.	
DR	MM;	111630; -	
DR	GO;	GO:0005887; C:Integral to plasma membrane; TAS.	
DR	InterPro;	IPRO01505; Ammonium transp.	
DR	InterPro;	IPRO02229; RheusRHD.	
DR	Pfam;	PF00909; Ammonium transp. 1.	
DR	PRINTS;	PRC0342; RHESUSRHD.	
KW	Erythrocyte; Transmembrane; Blood group antigen; Alternative splicing; Polymorphism.		
FT	INIT_MET	0	
FT	TRANSMEM	11	31
FT	TRANSMEM	43	63
FT	TRANSMEM	76	96
FT	TRANSMEM	124	144
FT	TRANSMEM	171	191
FT	TRANSMEM	202	222
FT	TRANSMEM	237	257
FT	TRANSMEM	264	284
FT	TRANSMEM	286	306
FT	TRANSMEM	330	350
FT	TRANSMEM	357	377
FT	VARSPLIC	162	312
FT	VARSPLIC	163	267
FT	VARSPLIC	313	353
FT	VARSPLIC		
FT	VARSPLIC	354	416
FT	VARSPLIC	358	416
FT	VARIANT	15	15
FT	VARIANT	35	35
FT	VARIANT	40	40
FT	VARIANT	59	59
FT	VARIANT		
FT	/FTID=VAR_006911.		
FT	/FTID=VAR_006912.		
FT	/FTID=VAR_006913.		
FT	/FTID=VAR_006914.		
FT	/FTID=VAR_006915.		
FT	/FTID=VAR_006916.		
FT	/FTID=VAR_006917.		
FT	/FTID=VAR_006918.		
FT	/FTID=VAR_006919.		
FT	/FTID=VAR_006920.		
FT	/FTID=VAR_006921.		
FT	/FTID=VAR_006922.		
FT	/FTID=VAR_006923.		
FT	/FTID=VAR_006924.		
FT	/FTID=VAR_006925.		
FT	/FTID=VAR_006926.		
FT	/FTID=VAR_006927.		
FT	/FTID=VAR_006928.		
FT	/FTID=VAR_006929.		
FT	/FTID=VAR_006930.		
FT	/FTID=VAR_006931.		
FT	/FTID=VAR_006932.		
FT	/FTID=VAR_006933.		
FT	/FTID=VAR_006934.		
FT	/FTID=VAR_006935.		
FT	/FTID=VAR_006936.		
FT	/FTID=VAR_006937.		
FT	/FTID=VAR_006938.		
FT	/FTID=VAR_006939.		
FT	/FTID=VAR_006940.		
FT	/FTID=VAR_006941.		
FT	/FTID=VAR_006942.		
FT	/FTID=VAR_006943.		
FT	/FTID=VAR_006944.		
FT	/FTID=VAR_006945.		
FT	/FTID=VAR_006946.		
FT	/FTID=VAR_006947.		
FT	/FTID=VAR_006948.		
FT	/FTID=VAR_006949.		
FT	/FTID=VAR_006950.		
FT	/FTID=VAR_006951.		
FT	/FTID=VAR_006952.		
FT	/FTID=VAR_006953.		
FT	/FTID=VAR_006954.		
FT	/FTID=VAR_006955.		
FT	/FTID=VAR_006956.		
FT	/FTID=VAR_006957.		
FT	/FTID=VAR_006958.		
FT	/FTID=VAR_006959.		
FT	/FTID=VAR_006960.		
FT	/FTID=VAR_006961.		
FT	/FTID=VAR_006962.		
FT	/FTID=VAR_006963.		
FT	/FTID=VAR_006964.		
FT	/FTID=VAR_006965.		
FT	/FTID=VAR_006966.		
FT	/FTID=VAR_006967.		
FT	/FTID=VAR_006968.		
FT	/FTID=VAR_006969.		
FT	/FTID=VAR_006970.		
FT	/FTID=VAR_006971.		
FT	/FTID=VAR_006972.		

```

OY 77 GSTEMVLVAVG-----VGAAPLGAIALTLTV--NFRHVFYAFSPFLHVNNPIARFYS 128
DB 237 GAVNTAAVVLVGDPRNIGAMWIGLISALVLSFPPFPFPMPIGAKAP----- 290
OY 129 VEPALIDEAVVTAAPAPGWSAMRLISMQAFHSYVWFGGLTGAVALIELIP-PEIKGLEFA 187
DB 291 --ATVDEARKLEERKSG-----SLVDFPKRPPCIFLRLL 323
OY 188 LGSIFVTLTLDSCRTKKQIPSLTLAAGS-----FITALVVIPOGL 228
DB 324 KMSLFLVVLAACTF-----SVYIAGLSIFLAKFLKXGYGTAAYANFLIGVNIPLAAL 378
OY 229 FALAI FLGLTLTRYPF 245
DB 379 G---MFGGILMRFRVF 392

RESULT 13
CLC4_HUMAN STANDARD; PRT; 760 AA.
AC P51793; Q9UBU1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chloride channel protein 4 (CLC-4).
GN CLC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA MEDLINE=94348498; PubMed=8069296;
RX van Slegtenhorst M.A., Bassi M.T., Borsani G., Wapenaar M.C.,
RA Ferrero G.B., de Conchitis L., Rugari E.I., Grillo A., Franco B.,
RA Zoghbi H.Y., Ballabio A.;
RT "A gene from the Xp22.3 region shares homology with voltage-gated
RT chloride channels."
RT Hum. Mol. Gene. 3:547-552 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=20035030; PubMed=10564087;
RA Kawasaki M., Fukuma T., Yamachi K., Sakamoto H., Marumo F.,
RA Sasaki S.;
RT "Identification of an acid-activated Cl- channel from human skeletal
RT muscles."
RT Am. J. Physiol. 277:C948-C954 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens epithelium;
RA Rae U.L.;
RT "A chloride channel (CLC-4) in human lens epithelium."
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transcellular transport.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Abundant in skeletal muscle and also
CC detectable in brain and heart.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; X77197; CAAS4417.1; -.
DR EMBL; AB019432; BAA77327.1; -.
DR EMBL; AF170492; AAD50981.1; -.
DR PIR; I37242; I37242.
DR Genew; HGNC:2022; CLC4.
DR MIM; 302910; -.
DR GO; GO:0005247; F: voltage-gated chloride channel activity; TAS.
DR GO; GO:0006810; P: transport; TAS.
DR InterPro; IPR00644; CBS domain.
DR InterPro; IPR01807; Cl-Channel_volt.
DR Pfam; PF00571; CBS; 2. Channel_volt.
DR Pfam; PF00654; CBS; 2. CLC; 1.
DR SMART; SM00116; CBS; 2.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Chloride channel; Chloride;
KW Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 68
FT TRANSMEM 88
FT TRANSMEM 148
FT TRANSMEM 168
FT TRANSMEM 202
FT TRANSMEM 222
FT TRANSMEM 259
FT TRANSMEM 279
FT TRANSMEM 330
FT TRANSMEM 350
FT TRANSMEM 365
FT TRANSMEM 385
FT TRANSMEM 442
FT TRANSMEM 462
FT TRANSMEM 467
FT TRANSMEM 487
FT TRANSMEM 508
FT TRANSMEM 528
FT TRANSMEM 535
FT TRANSMEM 555
FT TRANSMEM 599
FT TRANSMEM 659
FT TRANSMEM 697
FT TRANSMEM 747
FT TRANSMEM 178
FT TRANSMEM 198
FT TRANSMEM 498
FT TRANSMEM 499
FT TRANSMEM 659
FT TRANSMEM 659
SQ SEQUENCE 760 AA; 84916 MW; 3A5A25D1FEF3F217 CRC64;

Query Match 7.7%; Score 98; DB 1; Length 760;
Best Local Similarity 25.6%; Pred. No. 1.8;
Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

OY 80 EMLVIALVVG-----AAPLGAIALTLTVNFRHVFYAFSPFLHVNNPIARFYSVPLID 134
DB 254 EVLSAAAAAGVSVAFGAPDGV-----LFSLEEVSY--YPLKTL-----WSPFALV 300
OY 135 EAVVTAAPAPGWSAMRLISMQAFHSYVWFGGLTGAVALIELIPPEIKGLEFALC-SLFLV 193
DB 301 AAFILRSINPFGNS--RLVFLVYVEYHPWY-----MALPFIILGVFGGLMGTLLFI 350
OY 194 TLITDSCRTKK-----QIPSLTLAAGSFTALVVIIP 224
DB 351 RCNIAMCRRKRTTLGRYPVLEIVVTAITAIIVP 386

RESULT 14
RCCE_HUMAN STANDARD; PRT; 416 AA.
AC P18577; Q02163; Q02164; Q02165; Q16160; Q9UEC2; Q9UEC3;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Blood group Rh(CB) polypeptide (Rhesus C/E antigens) (RH30A) (RH19X)
DE (Rh polypeptide 1) (Rh1).
GN RHCE OR RHC OR RHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Bone marrow;
RX MEDLINE=90349591; PubMed=1696722;
RA Cherif-Zahar B., Bloy C., de Van Kim C., Blanchard D., Bailly P.,
RA Hermand P., Salmon C., Cartton J.-P., Colin Y.;
RT "Molecular cloning and protein structure of a human blood group Rh
RT polypeptide."
RT Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247 (1990).

```



```

CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; Z56277; CA93126.1; -.
CC EMBL; D50497; BAA09091.1; -.
CC InterPro; IPR000644; CBS domain.
CC InterPro; IPR001807; Cl-channel_volt.
CC Pfam; PF00571; CBS; 2.
CC Pfam; PF00654; voltage CIC; 1.
CC PRINTS; PR00762; CLCHANNEL.
CC SMART; SM00116; CBS; 2.
CC Voltage-gated channel; Transmembrane; CBS domain; Repeat.
CC Voltage-gated channel; Transmembrane; CBS domain; Repeat.
CC TRANSMEM 55 75 POTENTIAL.
CC TRANSMEM 133 153 POTENTIAL.
CC TRANSMEM 189 209 POTENTIAL.
CC TRANSMEM 246 266 POTENTIAL.
CC TRANSMEM 317 337 POTENTIAL.
CC TRANSMEM 352 372 POTENTIAL.
CC TRANSMEM 428 448 POTENTIAL.
CC TRANSMEM 453 473 POTENTIAL.
CC TRANSMEM 521 541 POTENTIAL.
CC TRANSMEM 585 614 CBS 1.
CC DOMAIN 682 733 CBS 2.
CC DOMAIN 315 315 H -> Y (IN REF. 2).
CC CONFLICT 746 AA; 83067 MW; 5F1D45F397003CE CRC64;
SQ SEQUENCE

Query Match 7.9%; Score 100.5; DB 1; Length 746;
Best Local Similarity 22.1%; Pred. No. 1.1; 68; Indels 75; Gaps 10;
Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;

27 YRRYEIAQGLKTSIAGLQWYPIGIAFGLLVIOGYEMWADPLFGLIFAGSTEMLVIAL 86
232 YRNEKREKREVLSSAAAG--VSVAFG----- 256
87 VCGAALGIALTLTLVNRHRYFASPEPLHVVQPIAFVSVPALLIDEAVAVTARAPAG 146
257 ---AIGGV---LFSLEVSY--YFPLKTL---WSFFPALVAALFTLRISINPFG 299
147 WSAWRILSMQIAFHSYVWFGGLGVAIALIPFRIKGLFALC-SLFTVTLTDSCKTK- 204
300 NS--RLVLFVFEHFTM-----HLFELVPIVIGIRGALGALFIRTNIAWCRKXT 349
205 ---QIPSLIAGSTTIALVTPGQ-----ALFALLIFGLI 239
350 TOLGKTPVEVVLITAITAILAFNPBYTMSLSLSELPNDCGL 395

RESULT 12
S212_HUMAN STANDARD; PRT; 643 AA.
ID S212_HUMAN
AC Q92959;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 21 member 2 (Prostaglandin transporter) (PCT).
GN SLC21A2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96379664; PubMed=8787677;
RA Lu R., Kanai N., Bao Y., Schuster V.L.;
RT "Cloning, in vitro expression, and tissue distribution of a human
RT prostaglandin transporter cDNA(hPGT).";
RT J. Clin. Invest. 98:1142-1149(1996).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=98289612; PubMed=9618293;
RA Lu R., Schuster V.L.;
RT "Molecular cloning of the gene for the human prostaglandin transporter

```

```

RT hPGT: gene organization, promoter activity, and chromosomal
RT localization."
RU Biochem. Biophys. Res. Commun. 246:805-812(1998).
CC CC
CC -1- FUNCTION: May mediate the release of newly synthesized
CC prostaglandins from cells, the trans epithelial transport of
CC prostaglandins, and the clearance of prostaglandins from the
CC circulation. Transports PGD2, as well as PGE1, PGE2 and PGF2a.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE SLC21 FAMILY OF TRANSPORTERS.
CC CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.sib-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL; U70867; AAC09469.1; -.
CC EMBL; AF056732; AAC62004.1; JOINED.
CC EMBL; AF056719; AAC62004.1; JOINED.
CC EMBL; AF056720; AAC62004.1; JOINED.
CC EMBL; AF056721; AAC62004.1; JOINED.
CC EMBL; AF056722; AAC62004.1; JOINED.
CC EMBL; AF056723; AAC62004.1; JOINED.
CC EMBL; AF056724; AAC62004.1; JOINED.
CC EMBL; AF056725; AAC62004.1; JOINED.
CC EMBL; AF056726; AAC62004.1; JOINED.
CC EMBL; AF056727; AAC62004.1; JOINED.
CC EMBL; AF056728; AAC62004.1; JOINED.
CC EMBL; AF056729; AAC62004.1; JOINED.
CC EMBL; AF056730; AAC62004.1; JOINED.
CC EMBL; AF056731; AAC62004.1; JOINED.
CC EMBL; AF056732; AAC62004.1; JOINED.
CC EMBL; AF056733; AAC62004.1; JOINED.
CC EMBL; AF056734; AAC62004.1; JOINED.
CC EMBL; AF056735; AAC62004.1; JOINED.
CC EMBL; AF056736; AAC62004.1; JOINED.
CC EMBL; AF056737; AAC62004.1; JOINED.
CC EMBL; AF056738; AAC62004.1; JOINED.
CC EMBL; AF056739; AAC62004.1; JOINED.
CC EMBL; AF056740; AAC62004.1; JOINED.
CC EMBL; AF056741; AAC62004.1; JOINED.
CC EMBL; AF056742; AAC62004.1; JOINED.
CC EMBL; AF056743; AAC62004.1; JOINED.
CC EMBL; AF056744; AAC62004.1; JOINED.
CC EMBL; AF056745; AAC62004.1; JOINED.
CC EMBL; AF056746; AAC62004.1; JOINED.
CC EMBL; AF056747; AAC62004.1; JOINED.
CC EMBL; AF056748; AAC62004.1; JOINED.
CC EMBL; AF056749; AAC62004.1; JOINED.
CC EMBL; AF056750; AAC62004.1; JOINED.
CC EMBL; AF056751; AAC62004.1; JOINED.
CC EMBL; AF056752; AAC62004.1; JOINED.
CC EMBL; AF056753; AAC62004.1; JOINED.
CC EMBL; AF056754; AAC62004.1; JOINED.
CC EMBL; AF056755; AAC62004.1; JOINED.
CC EMBL; AF056756; AAC62004.1; JOINED.
CC EMBL; AF056757; AAC62004.1; JOINED.
CC EMBL; AF056758; AAC62004.1; JOINED.
CC EMBL; AF056759; AAC62004.1; JOINED.
CC EMBL; AF056760; AAC62004.1; JOINED.
CC EMBL; AF056761; AAC62004.1; JOINED.
CC EMBL; AF056762; AAC62004.1; JOINED.
CC EMBL; AF056763; AAC62004.1; JOINED.
CC EMBL; AF056764; AAC62004.1; JOINED.
CC EMBL; AF056765; AAC62004.1; JOINED.
CC EMBL; AF056766; AAC62004.1; JOINED.
CC EMBL; AF056767; AAC62004.1; JOINED.
CC EMBL; AF056768; AAC62004.1; JOINED.
CC EMBL; AF056769; AAC62004.1; JOINED.
CC EMBL; AF056770; AAC62004.1; JOINED.
CC EMBL; AF056771; AAC62004.1; JOINED.
CC EMBL; AF056772; AAC62004.1; JOINED.
CC EMBL; AF056773; AAC62004.1; JOINED.
CC EMBL; AF056774; AAC62004.1; JOINED.
CC EMBL; AF056775; AAC62004.1; JOINED.
CC EMBL; AF056776; AAC62004.1; JOINED.
CC EMBL; AF056777; AAC62004.1; JOINED.
CC EMBL; AF056778; AAC62004.1; JOINED.
CC EMBL; AF056779; AAC62004.1; JOINED.
CC EMBL; AF056780; AAC62004.1; JOINED.
CC EMBL; AF056781; AAC62004.1; JOINED.
CC EMBL; AF056782; AAC62004.1; JOINED.
CC EMBL; AF056783; AAC62004.1; JOINED.
CC EMBL; AF056784; AAC62004.1; JOINED.
CC EMBL; AF056785; AAC62004.1; JOINED.
CC EMBL; AF056786; AAC62004.1; JOINED.
CC EMBL; AF056787; AAC62004.1; JOINED.
CC EMBL; AF056788; AAC62004.1; JOINED.
CC EMBL; AF056789; AAC62004.1; JOINED.
CC EMBL; AF056790; AAC62004.1; JOINED.
CC EMBL; AF056791; AAC62004.1; JOINED.
CC EMBL; AF056792; AAC62004.1; JOINED.
CC EMBL; AF056793; AAC62004.1; JOINED.
CC EMBL; AF056794; AAC62004.1; JOINED.
CC EMBL; AF056795; AAC62004.1; JOINED.
CC EMBL; AF056796; AAC62004.1; JOINED.
CC EMBL; AF056797; AAC62004.1; JOINED.
CC EMBL; AF056798; AAC62004.1; JOINED.
CC EMBL; AF056799; AAC62004.1; JOINED.
CC EMBL; AF056800; AAC62004.1; JOINED.
CC EMBL; AF056801; AAC62004.1; JOINED.
CC EMBL; AF056802; AAC62004.1; JOINED.
CC EMBL; AF056803; AAC62004.1; JOINED.
CC EMBL; AF056804; AAC62004.1; JOINED.
CC EMBL; AF056805; AAC62004.1; JOINED.
CC EMBL; AF056806; AAC62004.1; JOINED.
CC EMBL; AF056807; AAC62004.1; JOINED.
CC EMBL; AF056808; AAC62004.1; JOINED.
CC EMBL; AF056809; AAC62004.1; JOINED.
CC EMBL; AF056810; AAC62004.1; JOINED.
CC EMBL; AF056811; AAC62004.1; JOINED.
CC EMBL; AF056812; AAC62004.1; JOINED.
CC EMBL; AF056813; AAC62004.1; JOINED.
CC EMBL; AF056814; AAC62004.1; JOINED.
CC EMBL; AF056815; AAC62004.1; JOINED.
CC EMBL; AF056816; AAC62004.1; JOINED.
CC EMBL; AF056817; AAC62004.1; JOINED.
CC EMBL; AF056818; AAC62004.1; JOINED.
CC EMBL; AF056819; AAC62004.1; JOINED.
CC EMBL; AF056820; AAC62004.1; JOINED.
CC EMBL; AF056821; AAC62004.1; JOINED.
CC EMBL; AF056822; AAC62004.1; JOINED.
CC EMBL; AF056823; AAC62004.1; JOINED.
CC EMBL; AF056824; AAC62004.1; JOINED.
CC EMBL; AF056825; AAC62004.1; JOINED.
CC EMBL; AF056826; AAC62004.1; JOINED.
CC EMBL; AF056827; AAC62004.1; JOINED.
CC EMBL; AF056828; AAC62004.1; JOINED.
CC EMBL; AF056829; AAC62004.1; JOINED.
CC EMBL; AF056830; AAC62004.1; JOINED.
CC EMBL; AF056831; AAC62004.1; JOINED.
CC EMBL; AF056832; AAC62004.1; JOINED.
CC EMBL; AF056833; AAC62004.1; JOINED.
CC EMBL; AF056834; AAC62004.1; JOINED.
CC EMBL; AF056835; AAC62004.1; JOINED.
CC EMBL; AF056836; AAC62004.1; JOINED.
CC EMBL; AF056837; AAC62004.1; JOINED.
CC EMBL; AF056838; AAC62004.1; JOINED.
CC EMBL; AF056839; AAC62004.1; JOINED.
CC EMBL; AF056840; AAC62004.1; JOINED.
CC EMBL; AF056841; AAC62004.1; JOINED.
CC EMBL; AF056842; AAC62004.1; JOINED.
CC EMBL; AF056843; AAC62004.1; JOINED.
CC EMBL; AF056844; AAC62004.1; JOINED.
CC EMBL; AF056845; AAC62004.1; JOINED.
CC EMBL; AF056846; AAC62004.1; JOINED.
CC EMBL; AF056847; AAC62004.1; JOINED.
CC EMBL; AF056848; AAC62004.1; JOINED.
CC EMBL; AF056849; AAC62004.1; JOINED.
CC EMBL; AF056850; AAC62004.1; JOINED.
CC EMBL; AF056851; AAC62004.1; JOINED.
CC EMBL; AF056852; AAC62004.1; JOINED.
CC EMBL; AF056853; AAC62004.1; JOINED.
CC EMBL; AF056854; AAC62004.1; JOINED.
CC EMBL; AF056855; AAC62004.1; JOINED.
CC EMBL; AF056856; AAC62004.1; JOINED.
CC EMBL; AF056857; AAC62004.1; JOINED.
CC EMBL; AF056858; AAC62004.1; JOINED.
CC EMBL; AF056859; AAC62004.1; JOINED.
CC EMBL; AF056860; AAC62004.1; JOINED.
CC EMBL; AF056861; AAC62004.1; JOINED.
CC EMBL; AF056862; AAC62004.1; JOINED.
CC EMBL; AF056863; AAC62004.1; JOINED.
CC EMBL; AF056864; AAC62004.1; JOINED.
CC EMBL; AF056865; AAC62004.1; JOINED.
CC EMBL; AF056866; AAC62004.1; JOINED.
CC EMBL; AF056867; AAC62004.1; JOINED.
CC EMBL; AF056868; AAC62004.1; JOINED.
CC EMBL; AF056869; AAC62004.1; JOINED.
CC EMBL; AF056870; AAC62004.1; JOINED.
CC EMBL; AF056871; AAC62004.1; JOINED.
CC EMBL; AF056872; AAC62004.1; JOINED.
CC EMBL; AF056873; AAC62004.1; JOINED.
CC EMBL; AF056874; AAC62004.1; JOINED.
CC EMBL; AF056875; AAC62004.1; JOINED.
CC EMBL; AF056876; AAC62004.1; JOINED.
CC EMBL; AF056877; AAC62004.1; JOINED.
CC EMBL; AF056878; AAC62004.1; JOINED.
CC EMBL; AF056879; AAC62004.1; JOINED.
CC EMBL; AF056880; AAC62004.1; JOINED.
CC EMBL; AF056881; AAC62004.1; JOINED.
CC EMBL; AF056882; AAC62004.1; JOINED.
CC EMBL; AF056883; AAC62004.1; JOINED.
CC EMBL; AF056884; AAC62004.1; JOINED.
CC EMBL; AF056885; AAC62004.1; JOINED.
CC EMBL; AF056886; AAC62004.1; JOINED.
CC EMBL; AF056887; AAC62004.1; JOINED.
CC EMBL; AF056888; AAC62004.1; JOINED.
CC EMBL; AF056889; AAC62004.1; JOINED.
CC EMBL; AF056890; AAC62004.1; JOINED.
CC EMBL; AF056891; AAC62004.1; JOINED.
CC EMBL; AF056892; AAC62004.1; JOINED.
CC EMBL; AF056893; AAC62004.1; JOINED.
CC EMBL; AF056894; AAC62004.1; JOINED.
CC EMBL; AF056895; AAC62004.1; JOINED.
CC EMBL; AF056896; AAC62004.1; JOINED.
CC EMBL; AF056897; AAC62004.1; JOINED.
CC EMBL; AF056898; AAC62004.1; JOINED.
CC EMBL; AF056899; AAC62004.1; JOINED.
CC EMBL; AF056900; AAC62004.1; JOINED.
CC EMBL; AF056901; AAC62004.1; JOINED.
CC EMBL; AF056902; AAC62004.1; JOINED.
CC EMBL; AF056903; AAC62004.1; JOINED.
CC EMBL; AF056904; AAC62004.1; JOINED.
CC EMBL; AF056905; AAC62004.1; JOINED.
CC EMBL; AF056906; AAC62004.1; JOINED.
CC EMBL; AF056907; AAC62004.1; JOINED.
CC EMBL; AF056908; AAC62004.1; JOINED.
CC EMBL; AF056909; AAC62004.1; JOINED.
CC EMBL; AF056910; AAC62004.1; JOINED.
CC EMBL; AF056911; AAC62004.1; JOINED.
CC EMBL; AF056912; AAC62004.1; JOINED.
CC EMBL; AF056913; AAC62004.1; JOINED.
CC EMBL; AF056914; AAC62004.1; JOINED.
CC EMBL; AF056915; AAC62004.1; JOINED.
CC EMBL; AF056916; AAC62004.1; JOINED.
CC EMBL; AF056917; AAC62004.1; JOINED.
CC EMBL; AF056918; AAC62004.1; JOINED.
CC EMBL; AF056919; AAC62004.1; JOINED.
CC EMBL; AF056920; AAC62004.1; JOINED.
CC EMBL; AF056921; AAC62004.1; JOINED.
CC EMBL; AF056922; AAC62004.1; JOINED.
CC EMBL; AF056923; AAC62004.1; JOINED.
CC EMBL; AF056924; AAC62004.1; JOINED.
CC EMBL; AF056925; AAC62004.1; JOINED.
CC EMBL; AF056926; AAC62004.1; JOINED.
CC EMBL; AF056927; AAC62004.1; JOINED.
CC EMBL; AF056928; AAC62004.1; JOINED.
CC EMBL; AF056929; AAC62004.1; JOINED.
CC EMBL; AF056930; AAC62004.1; JOINED.
CC EMBL; AF056931; AAC62004.1; JOINED.
CC EMBL; AF056932; AAC62004.1; JOINED.
CC EMBL; AF056933; AAC62004.1; JOINED.
CC EMBL; AF056934; AAC62004.1; JOINED.
CC EMBL; AF056935; AAC62004.1; JOINED.
CC EMBL; AF056936; AAC62004.1; JOINED.
CC EMBL; AF056937; AAC62004.1; JOINED.
CC EMBL; AF056938; AAC62004.1; JOINED.
CC EMBL; AF056939; AAC62004.1; JOINED.
CC EMBL; AF056940; AAC62004.1; JOINED.
CC EMBL; AF056941; AAC62004.1; JOINED.
CC EMBL; AF056942; AAC62004.1; JOINED.
CC EMBL; AF056943; AAC62004.1; JOINED.
CC EMBL; AF056944; AAC62004.1; JOINED.
CC EMBL; AF056945; AAC62004.1; JOINED.
CC EMBL; AF056946; AAC62004.1; JOINED.
CC EMBL; AF056947; AAC62004.1; JOINED.
CC EMBL; AF056948; AAC62004.1; JOINED.
CC EMBL; AF056949; AAC62004.1; JOINED.
CC EMBL; AF056950; AAC62004.1; JOINED.
CC EMBL; AF056951; AAC62004.1; JOINED.
CC EMBL; AF056952; AAC62004.1; JOINED.
CC EMBL; AF056953; AAC62004.1; JOINED.
CC EMBL; AF056954; AAC62004.1; JOINED.
CC EMBL; AF056955; AAC62004.1; JOINED.
CC EMBL; AF056956; AAC62004.1; JOINED.
CC EMBL; AF056957; AAC62004.1; JOINED.
CC EMBL; AF056958; AAC62004.1; JOINED.
CC EMBL; AF056959; AAC62004.1; JOINED.
CC EMBL; AF056960; AAC62004.1; JOINED.
CC EMBL; AF056961; AAC62004.1; JOINED.
CC EMBL; AF056962; AAC62004.1; JOINED.
CC EMBL; AF056963; AAC62004.1; JOINED.
CC EMBL; AF056964; AAC62004.1; JOINED.
CC EMBL; AF056965; AAC62004.1; JOINED.
CC EMBL; AF056966; AAC62004.1; JOINED.
CC EMBL; AF056967; AAC62004.1; JOINED.
CC EMBL; AF056968; AAC62004.1; JOINED.
CC EMBL; AF056969; AAC62004.1; JOINED.
CC EMBL; AF056970; AAC62004.1; JOINED.
CC EMBL; AF056971; AAC62004.1; JOINED.
CC EMBL; AF056972; AAC62004.1; JOINED.
CC EMBL; AF056973; AAC62004.1; JOINED.
CC EMBL; AF056974; AAC62004.1; JOINED.
CC EMBL; AF056975; AAC62004.1; JOINED.
CC EMBL; AF056976; AAC62004.1; JOINED.
CC EMBL; AF056977; AAC62004.1; JOINED.
CC EMBL; AF056978; AAC62004.1; JOINED.
CC EMBL; AF056979; AAC62004.1; JOINED.
CC EMBL; AF056980; AAC62004.1; JOINED.
CC EMBL; AF056981; AAC62004.1; JOINED.
CC EMBL; AF056982; AAC62004.1; JOINED.
CC EMBL; AF056983; AAC62004.1; JOINED.
CC EMBL; AF056984; AAC62004.1; JOINED.
CC EMBL; AF056985; AAC62004.1; JOINED.
CC EMBL; AF056986; AAC62004.1; JOINED.
CC EMBL; AF056987; AAC62004.1; JOINED.
CC EMBL; AF056988; AAC62004.1; JOINED.
CC EMBL; AF056989; AAC62004.1; JOINED.
CC EMBL; AF056990; AAC62004.1; JOINED.
CC EMBL; AF056991; AAC62004.1; JOINED.
CC EMBL; AF056992; AAC62004.1; JOINED.
CC EMBL; AF056993; AAC62004.1; JOINED.
CC EMBL; AF056994; AAC62004.1; JOINED.
CC EMBL; AF056995; AAC62004.1; JOINED.
CC EMBL; AF056996; AAC62004.1; JOINED.
CC EMBL; AF056997; AAC62004.1; JOINED.
CC EMBL; AF056998; AAC62004.1; JOINED.
CC EMBL; AF056999; AAC62004.1; JOINED.
CC EMBL; AF057000; AAC62004.1; JOINED.
CC EMBL; AF057001; AAC62004.1; JOINED.
CC EMBL; AF057002; AAC62004.1; JOINED.
CC EMBL; AF057003; AAC62004.1; JOINED.
CC EMBL; AF057004; AAC62004.1; JOINED.
CC EMBL; AF057005; AAC62004.1; JOINED.
CC EMBL; AF057006; AAC62004.1; JOINED.
CC EMBL; AF057007; AAC62004.1; JOINED.
CC EMBL; AF057008; AAC62004.1; JOINED.
CC EMBL; AF057009; AAC62004.1; JOINED.
CC EMBL; AF057010; AAC62004.1; JOINED.
CC EMBL; AF057011; AAC62004.1; JOINED.
CC EMBL; AF057012; AAC62004.1; JOINED.
CC EMBL; AF057013; AAC62004.1; JOINED.
CC EMBL; AF057014; AAC62004.1; JOINED.
CC EMBL; AF057015; AAC62004.1; JOINED.
CC EMBL; AF057016; AAC62004.1; JOINED.
CC EMBL; AF057017; AAC62004.1; JOINED.
CC EMBL; AF057018; AAC62004.1; JOINED.
CC EMBL; AF057019; AAC62004.1; JOINED.
CC EMBL; AF057020; AAC62004.1; JOINED.
CC EMBL; AF057021; AAC62004.1; JOINED.
CC EMBL; AF057022; AAC62004.1; JOINED.
CC EMBL; AF057023; AAC62004.1; JOINED.
CC EMBL; AF057024; AAC62004.1; JOINED.
CC EMBL; AF057025; AAC62004.1; JOINED.
CC EMBL; AF057026; AAC62004.1; JOINED.
CC EMBL; AF057027; AAC62004.1; JOINED.
CC EMBL; AF057028; AAC62004.1; JOINED.
CC EMBL; AF057029; AAC62004.1; JOINED.
CC EMBL; AF057030; AAC62004.1; JOINED.
CC EMBL; AF057031; AAC62004.1; JOINED.
CC EMBL; AF057032; AAC62004.1; JOINED.
CC EMBL; AF057033; AAC62004.1; JOINED.
CC EMBL; AF057034; AAC62004.1; JOINED.
CC EMBL; AF057035; AAC62004.1; JOINED.
CC EMBL; AF057036; AAC62004.1; JOINED.
CC EMBL; AF057037; AAC62004.1; JOINED.
CC EMBL; AF057038; AAC62004.1; JOINED.
CC EMBL; AF057039; AAC62004.1; JOINED.
CC EMBL; AF057040; AAC62004.1; JOINED.
CC EMBL; AF057041; AAC62004.1; JOINED.
CC EMBL; AF057042; AAC62004.1; JOINED.
CC EMBL; AF057043; AAC62004.1; JOINED.
CC EMBL; AF057044; AAC62004.1; JOINED.
CC EMBL; AF057045; AAC62004.1; JOINED.
CC EMBL; AF057046; AAC62004.1; JOINED.
CC EMBL; AF057047; AAC62004.1; JOINED.
CC EMBL; AF057048; AAC62004.1; JOINED.
CC EMBL; AF057049; AAC62004.1; JOINED.
CC EMBL; AF057050; AAC62004.1; JOINED.
CC EMBL; AF057051; AAC62004.1; JOINED.
CC EMBL; AF057052; AAC62004.1; JOINED.
CC EMBL; AF057053; AAC62004.1; JOINED.
CC EMBL; AF057054; AAC62004.1; JOINED.
CC EMBL; AF057055; AAC62004.1; JOINED.
CC EMBL; AF057056; AAC62004.1; JOINED.
CC EMBL; AF057057; AAC62004.1; JOINED.
CC EMBL; AF057058; AAC62004.1; JOINED.
CC EMBL; AF057059; AAC62004.1; JOINED.
CC EMBL; AF057060; AAC62004.1; JOINED.
CC EMBL; AF057061; AAC62004.1; JOINED.
CC EMBL; AF057062; AAC62004.1; JOINED.
CC EMBL; AF057063; AAC62004.1; JOINED.
CC EMBL; AF057064; AAC62004.1; JOINED.
CC EMBL; AF057065; AAC62004.1; JOINED.
CC EMBL; AF057066; AAC62004.1; JOINED.
CC EMBL; AF057067; AAC62004.1; JOINED.
CC EMBL; AF057068; AAC62004.1; JOINED.
CC EMBL; AF057069; AAC62004.1; JOINED.
CC EMBL; AF057070; AAC62004.1; JOINED.
CC EMBL; AF057071; AAC62004.1; JOINED.
CC EMBL; AF057072; AAC62004.1; JOINED.
CC EMBL; AF057073; AAC62004.1; JOINED.
CC EMBL; AF057074; AAC62004.1; JOINED.
CC EMBL; AF057075; AAC62004.1; JOINED.
CC EMBL; AF057076; AAC62004.1; JOINED.
CC EMBL; AF057077; AAC62004.1; JOINED.
CC EMBL; AF057078; AAC62004.1; JOINED.
CC EMBL; AF057079; AAC62004.1; JOINED.
CC EMBL; AF057080; AAC62004.1; JOINED.
CC EMBL; AF057081; AAC62004.1; JOINED.
CC EMBL; AF057082; AAC62004.1; JOINED.
CC EMBL; AF057083; AAC62004.1; JOINED.
CC EMBL; AF057084; AAC62004.1; JOINED.
CC EMBL; AF057085; AAC62004.1; JOINED.
CC EMBL; AF057086; AAC62004.1; JOINED.
CC EMBL; AF057087; AAC62004.1; JOINED.
CC EMBL; AF057088; AAC62004.1; JOINED.
CC EMBL; AF057089; AAC62004.1; JOINED.
CC EMBL; AF057090; AAC62004.1; JOINED.
CC EMBL; AF057091; AAC62004.1; JOINED.
CC EMBL; AF057092; AAC62004.1; JOINED.
CC EMBL; AF057093; AAC62004.1; JOINED.
CC EMBL; AF057094; AAC62004.1; JOINED.
CC EMBL; AF057095; AAC62004.1; JOINED.
CC EMBL; AF057096; AAC62004.1; JOINED.
CC EMBL; AF057097; AAC62004.1; JOINED.
CC EMBL; AF057098; AAC62004.1; JOINED.
CC EMBL; AF057099; AAC62004.1; JOINED.
CC EMBL; AF057100; AAC62004.1; JOINED.
CC EMBL; AF057101; AAC62004.1; JOINED.
CC EMBL; AF057102; AAC62004.1; JOINED.
CC EMBL; AF057103; AAC62004.1; JOINED.
CC EMBL; AF057104; AAC62004.1; JOINED.
CC EMBL; AF057105; AAC62004.1; JOINED.
CC EMBL; AF057106; AAC62004.1; JOINED.
CC EMBL; AF057107; AAC62004.1; JOINED.
CC EMBL; AF057108; AAC62004.1; JOINED.
CC EMBL; AF057109; AAC62004.1; JOINED.
CC EMBL; AF057110; AAC62004.1; JOINED.
CC EMBL; AF057111; AAC62004.1; JOINED.
CC EMBL; AF057112; AAC62004.1; JOINED.
CC EMBL; AF057113; AAC62004.1; JOINED.
CC EMBL; AF057114; AAC62004.1; JOINED.
CC EMBL; AF057115; AAC62004.1; JOINED.
CC EMBL; AF057116; AAC62004.1; JOINED.
CC EMBL; AF057117; AAC62004.1; JOINED.
CC EMBL; AF057118; AAC62004.1; JOINED.
CC EMBL; AF057119; AAC62004.1; JOINED.
CC EMBL; AF057120; AAC62004.1; JOINED.
CC EMBL; AF057121; AAC62004.1; JOINED.
CC EMBL; AF057122; AAC62004.1; JOINED.
CC EMBL; AF057123; AAC62004.1; JOINED.
CC EMBL; AF057124; AAC62004.1; JOINED.
CC EMBL; AF057125; AAC62004.1; JOINED.
CC EMBL; AF057126; AAC62004.1; JOINED.
CC EMBL; AF057127; AAC62004.1; JOINED.
CC EMBL; AF057128; AAC62004
```

```

Db      232 YRKNAKREVLASAAAAG--VSAVFG-----256
Qy      87 VGGAAPLGAIALTTLVNFRHVFYAFSPFLHVKNPIAFYSVAFALIDEAVATAAPAG 146
Db      257 ----APIGV-----LFSLEVSY--YFPLKTL-----WRSFFALVAAFTLSINPFG 299
Qy      147 MSAMRLISMQIAFHSYVFGSLTGVAIAELIPEIKGLEFALC-SLFTVTLTDSGCRTK- 204
Db      300 NS--RLVLFYVEFHTPW-----HLFELVPEIVLIGFGGLMGALFIRTNIAWCRKRT 349
Qy      205 ----QIPSLLAGISFTIALVVIPOQ-----ALFALLIFGLL 239
Db      350 TOLGKYPVEVLIVTAITAILAFNPEYTRMSTSELISELFDGGL 395

RESULT 10
CLC5_MOUSE STANDARD; PRT; 746 AA.
ID CLC5_MOUSE
AC 09WDA4:
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLC5 OR CLC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=99303559; PubMed=10373326;
RA Tanka K., Fisher S.E., Craig I.W.;
RT "Characterization of novel promoter and enhancer elements of the mouse
RT homologue of the Dent disease gene, CLC5, implicated in X-linked
RT hereditary nephrolithiasis.";
RL Genomics 58:281-292(1999).
CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transepithelial transport. May play an important role in renal
CC tubular function.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Kidney specific.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF134117; ABD28473.1; -.
DR MGD; WGI:99486; CLC5.
DR InterPro; IPR000644; CBS_domain.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CICHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Chloride channel; Chloride;
KW Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 55
FT TRANSMEM 75
FT TRANSMEM 133
FT TRANSMEM 153
FT TRANSMEM 189
FT TRANSMEM 209
FT TRANSMEM 246
FT TRANSMEM 266
FT TRANSMEM 317
FT TRANSMEM 337
FT TRANSMEM 352
FT TRANSMEM 372
FT TRANSMEM 428
FT TRANSMEM 448
FT TRANSMEM POTENTIAL.

```

```

FT TRANSMEM 453
FT TRANSMEM 473
FT TRANSMEM 494
FT TRANSMEM 514
FT TRANSMEM 521
FT TRANSMEM 541
FT TRANSMEM 582
FT TRANSMEM 614
FT TRANSMEM 682
FT TRANSMEM 733
FT TRANSMEM 83100 MW; DBFJAEAFQ331A08 CRC64;
SQ SEQUENCE 746 AA; 83100 MW; DBFJAEAFQ331A08 CRC64;

Query Match 7.9%; Score 100.5; DB 1; Length 746;
Best Local Similarity 22.1%; Pred. No. 1.1;
Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;

Qy      27 YRRYEIAQGLKTSIAGLGMVPIGIAFGLLVIOGYEMAAPLPSGLIFAGSTEMLVIAL 86
Db      232 YRKNAKREVLASAAAAG--VSAVFG-----256
Qy      87 VGGAAPLGAIALTTLVNFRHVFYAFSPFLHVKNPIAFYSVAFALIDEAVATAAPAG 146
Db      257 ----APIGV-----LFSLEVSY--YFPLKTL-----WRSFFALVAAFTLSINPFG 299
Qy      147 MSAMRLISMQIAFHSYVFGSLTGVAIAELIPEIKGLEFALC-SLFTVTLTDSGCRTK- 204
Db      300 NS--RLVLFYVEFHTPW-----HLFELVPEIVLIGFGGLMGALFIRTNIAWCRKRT 349
Qy      205 ----QIPSLLAGISFTIALVVIPOQ-----ALFALLIFGLL 239
Db      350 TOLGKYPVEVLIVTAITAILAFNPEYTRMSTSELISELFDGGL 395

RESULT 11
CLC5_RAT STANDARD; PRT; 746 AA.
ID CLC5_RAT
AC P51796; P70642;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Chloride channel protein 5 (CLC-5).
GN CLC5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96125100; PubMed=8537381;
RA Steinmeyer K., Schwappach B., Bens M., Vandewalle A., Jentsch T.U.;
RT "Cloning and functional expression of rat CLC-5, a chloride channel
RT related to kidney disease.";
RL J. Biol. Chem. 270:31172-31177(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96215316; PubMed=8626585;
RA Sakamoto H., Kawasaki M., Uchida S., Sasaki S., Maruno F.;
RT "Identification of a new outwardly rectifying Cl- channel that
RT belongs to a subfamily of the CLC Cl- channels.";
RL J. Biol. Chem. 271:10210-10216(1996).
CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transepithelial transport. May play an important role in renal
CC tubular function.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Kidney specific.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CN CLCN5 OR CLCN2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=96121370; PubMed=8575751;
 RA Fisher S.E., van Bavel I., Lloyd S.E., Pearce S.H.S.,
 RA Thakker R.V., Craig I.W.;
 RT "Cloning and characterization of CLCN5, the human kidney chloride
 channel gene implicated in Dent disease (an X-linked hereditary
 nephrolithiasis)." J.
 RL Hum. Mol. Genet. 29:598-606(1995).
 RN [2]
 RP SEQUENCE OF 487-746 FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95179126; PubMed=7874126;
 RA Fisher S., Black G.C.M., Lloyd S.E., Hatchwell E., Wrong O.,
 RA Thakker R.V., Craig I.W.;
 RT "Isolation and partial characterization of a chloride channel gene
 which is expressed in kidney and is a candidate for Dent's disease
 (an X-linked hereditary nephrolithiasis)." J.
 RL Hum. Mol. Genet. 3:2053-2059(1994).
 RN [3]
 RP TISSUE SPECIFICITY.
 RC TISSUE=Arteric endothelium, and Vascular smooth muscle;
 RX MEDLINE=99222487; PubMed=10198195;
 RA Lamb P.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
 RA Schulte B.C.;
 RT "Expression of CLCN voltage-gated chloride channel genes in human
 blood vessels." J.
 RL J. Mol. Cell. Cardiol. 31:657-666(1999)..
 RN [4]
 RP VARIANTS NPHL ARG-200; GLU-506 AND PRO-520, AND VARIANT XLRH LEU-244.
 RX MEDLINE=96158876; PubMed=8559248;
 RA Lloyd S.E., Pearce S.H.S., Fisher S.E., Steimeyer K., Schwappach B.,
 RA Scheinman S.J., Harding B., Bolino A., Devoto M., Goodyer P.,
 RA Ridsden S.P.A., Wrong O., Jentsch T.J., Craig I.W., Thakker R.V.;
 RT "A common molecular basis for three inherited kidney stone diseases." J.
 RL Nature 379:445-449(1996).
 RN [5]
 RP VARIANTS NPHL.
 RX MEDLINE=97402204; PubMed=9259268;
 RA Lloyd S.E., Guenther W., Pearce S.H.S., Thomson A., Bianchi M.L.,
 RA Bosio M., Craig I.W., Fisher S.E., Scheinman S.J., Wrong O.,
 RA Jentsch T.J., Thakker R.V.;
 RT "Characterization of renal chloride channel, CLCN5, mutations in
 hypercalcaemic nephrolithiasis (kidney stones) disorders." J.
 RL Hum. Mol. Genet. 6:1233-1239(1997).
 CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
 several functions including the regulation of cell volume,
 membrane potential stabilization, signal transduction and
 transepithelial transport. May play an important role in renal
 tubular function.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Kidney. Moderately expressed in aortic
 vascular smooth muscle and endothelial cells, and at a slightly
 higher level in the coronary vascular smooth muscle.
 CC -1- DISEASE: Defects in CLCN5 are a cause of X-linked recessive
 hypophosphatemic rickets (XLRH) [MIM:300008].
 CC -1- DISEASE: Defects in CLCN5 are the cause of idiopathic low
 molecular weight proteinuria of Japanese children (JILP)
 [MIM:300008].
 CC -1- DISEASE: Defects in CLCN5 are the cause of X-linked recessive
 nephrolithiasis type 2 (NPHL2) [MIM:300009], also known as Dent
 disease. NPHL2 is a renal form of Fanconi syndrome, characterized
 by hypercalcaemia, nephrocalcinosis, renal stones and progressive
 renal failure.
 CC -1- DISEASE: Defects in CLCN5 are the cause of X-linked recessive
 nephrolithiasis type 1 (NPHL1) [MIM:310468]; also designated XRN.
 CC NPHL1 is characterized by hypercalcaemia, nephrocalcinosis, renal

CC stones and renal insufficiency, but differs from Dent disease in
 CC that the patients lack urinary acidification defects, rickets, and
 CC osteomalacia.
 CC -1- SIMILARITY: Contains 2 CBS domains.
 CC -1- SIMILARITY: Contains 2 CBS domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91906; CA63000.1; -;
 DR EMBL; X81836; CA57430.1; -;
 DR PIR; I37277; I37277.
 DR GeneW; HGNC:2023; CLCN5.
 DR MIM; 300008; -;
 DR MIM; 300009; -;
 DR MIM; 310468; -;
 DR GO; GO:0005887; C:Integral to plasma membrane; TAS.
 DR GO; GO:0005254; F:chloride channel activity; TAS.
 DR GO; GO:0007586; P:excretion; TAS.
 DR GO; GO:0006810; P:transport; TAS.
 DR InterPro; IPR000644; CBS_domain.
 DR InterPro; IPR01807; Cl_Channel_volt.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM00116; CBS; 2.
 KW Ionic channel; Ion transport; Chloride channel; Chloride;
 KW Voltage-gated channel; Transmembrane; CBS domain; Repeat;
 KW Disease mutation.
 FT TRANSSEM 55 75 POTENTIAL.
 FT TRANSSEM 133 153 POTENTIAL.
 FT TRANSSEM 189 209 POTENTIAL.
 FT TRANSSEM 246 266 POTENTIAL.
 FT TRANSSEM 317 337 POTENTIAL.
 FT TRANSSEM 352 372 POTENTIAL.
 FT TRANSSEM 428 448 POTENTIAL.
 FT TRANSSEM 453 473 POTENTIAL.
 FT TRANSSEM 494 514 POTENTIAL.
 FT TRANSSEM 521 541 POTENTIAL.
 FT DOMAIN 585 645 CBS 1.
 FT DOMAIN 682 733 CBS 2.
 FT VARIANT 30 30 R -> RH (in NPHL2).
 FT VARIANT 57 57 /FtId=VAR_001615.
 FT VARIANT 200 200 L -> V (in NPHL2).
 FT VARIANT 244 244 L -> R (in NPHL2).
 FT VARIANT 280 280 S -> L (in XLRH).
 FT VARIANT 280 280 /FtId=VAR_001618.
 FT VARIANT 506 506 R -> P (in JILP).
 FT VARIANT 512 512 /FtId=VAR_001619.
 FT VARIANT 512 512 G -> E (in NPHL1).
 FT VARIANT 512 512 /FtId=VAR_001620.
 FT VARIANT 512 512 G -> R (in NPHL2) abolishes the chloride
 FT VARIANT 512 512 currents).
 FT VARIANT 520 520 /FtId=VAR_001621.
 FT VARIANT 520 520 S -> P (in NPHL2).
 FT VARIANT 527 527 /FtId=VAR_001622.
 FT VARIANT 527 527 E -> D (in NPHL2) abolishes the chloride
 FT VARIANT 527 527 currents; total loss of function).
 FT VARIANT 527 527 /FtId=VAR_001623.
 SQ SEQUENCE 746 AA; 83146 MW; EF913C5B40C85D8 CRC64;
 Query Match 8.0%; Score 101.5; DB 1; Length 746;
 Best Local Similarity 22.1%; Pred. No. 0.92;
 Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;
 27 YRRFVINGKLTSLAAGIGWYPIGIAFGLLVIVGYGYEMWMAPIFSGIIFAGSTEMVIAL 86

DE NAD(P)-H-gulonone oxidoreductase chain 4, chloroplast (EC 1.6.5.-)
DE NAD(P)-H dehydrogenase, chain 4) (NADH-plastoquinone oxidoreductase
DE chain 4).
DE NDH.
GN NDH.
OS Psilotum nudum (Whisk fern).
OC Chloroplast.
OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Monilliformes; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
OX NCBI_TaxID=3240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kinsyoku;
RA Wakasugi T., Nishikawa A., Yamada K., Sugiyura M.;
RT "Complete nucleotide sequence of the chloroplast genome from a fern,
CC Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P) (+) +
CC plastoquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>
CC or send an email to license@ebi.ac.uk).CC
CC
DR EMBL: AP004638; BAB64274.1; ALT_INIT.
DR HAMAP: MF_004631; -1
DR InterPro: IPR003918; NADHdb Oxred4.
DR InterPro: IPR001750; Oxidored_g1.
DR Pfam: PF00361; Oxidored_g1; 1.
DR PRINTS: PRO1437; NOOXDRDTASE4.
DR OXIDOREDUCTASE; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ SEQUENCE 498 AA; 55958 MW; 7C89E6003E8EAD95 CRC64;

Query Match 8.0%; Score 101.5; DB 1; Length 498;
Best Local Similarity 22.5%; Pred. NO. 0.64; Indels 117; Gaps 16;
Matches 69; Conservative 40; Mismatches

QY 44 LGMY-DIGIAFGLLVIOGYEW-----WAAP-LFSGLIAGSTEMLVIAL 86
DB 119 LAMWSQIGLFASONILFFEFMELELPVYLLISMGGKRLTSSTFLLHYTAGSGIFL 178
QY 87 VCGAPRG-----AIALTLTVNFRHFVAFSPFL-HVYNKPIARF 126
DB 179 LVGSLTNGLYGSGNSTFDQLITKRSYPIAVETLY-----FSFLIAYAVLPILPF 230
QY 127 -----YSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYVFG-----GL 168
DB 231 HTWLPDTRGEAHYSTCML-----AGILKMGGYGLIR-INMELPHAFPLPVLVTWGA 285
QY 169 TGVVIALEIPEIKGL--FALGSL-----FTLLTDSGR--TKQIPSLI----- 210
DB 286 VQIVYASLIQSFSLPIKRIAYSSVMGVIYIGISSTDTGNTGALIQMISHGILGAL 345
QY 211 -LAGLS-----FTTALVVPQALFAA-LTIFGLL 239
DB 346 FFLAGIYSDRTQTLTLQDGLGIAVPMRPLFTMGFFFSIASIALTPGMSFVAEFLIFIGIV 405
QY 240 TIRYEFLL 246
DB 406 TNPYISL 412

RESULT 9
CLCS_HUMAN STANDARD; PRT; 746 AA.
AC PS1795;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chloride channel protein 5 (CLIC-5).

SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirtpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000353; AAC75729.1; -
 CC PIR: C65048; C65048.
 CC EcoGene: EG13528; YGAZ.
 DR InterPro: IPR004471; Azic.
 DR Pfam: PF03591; Azic; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 25 45
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 110 130 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 245 26107 MW; 22ACB8ABD7D651B CRC64;
 SQ SEQUENCE 245 AA; 26107 MW; 22ACB8ABD7D651B CRC64;
 Query Match 15.5%; Score 197; DB 1; Length 245;
 Best Local Similarity 29.1%; Pred. No. 1,1e-08; Mismatches 95; Indels 30; Gaps 9;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;
 QY 34 QGLKTSLAAGLGMVPIGIAFGILVIGYEWMAAPFSGILFAGSTEMLVVGAAPL 93
 DB 17 EGCKDLPFIVISYIPAFAGLNTLGFSPLESVFSCIIYAGSQPITMLAAGSL 76
 QY 94 GAIALTLTLVNFPHVYAFSPFLAVV---KPIAFYVFPALIDBAVNTAR---PA 145
 DB 77 WIALTVMAVDVHVYVGSLSRSRIIQLQSKRTALW--AFGILTEVFPAATAKVRNR 134
 QY 146 GMSAMFLISMQIAFHSVWVF---GGILGVAAIAELIPFIKGLFALCSLFTLLTDSQR 201
 DB 135 RSEBNMMIGIAFSMSWVGTGVIAGFSGSLGQVPAVEALGLFPLPALFMSFLIASPQ 194
 QY 202 TKKQISLLAIGLSFTIALV-VIPQALPA--ALI--IFLGLLT--IYFPLG 247
 DB 195 RKQ-----SLCVTAALVGAAGVTLFSPVAILGIVCGCTIALIQAFWQG 240
 RESULT 6
 YH55_ARCFU STANDARD; PRT; 219 AA.
 AC 028519;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1755.
 GN AF1755.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 ON NCBI_Taxid=2234;
 RX STRAIN=VC-16 / DSM 4304 / ATCC 49556;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

Retchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Keriavage A.R., Graham D.E., Kyriides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Arlanch P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000982; AAB89495.1; -
 CC PIR: B69469; B69469.
 DR TIGR: AF1755; -
 DR InterPro: IPR004471; Azic.
 DR Pfam: PF03591; Azic; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 37 57 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT TRANSMEM 219 2337 MW; 265A2C88DA93EBB6 CRC64;
 SQ SEQUENCE 219 AA; 2337 MW; 265A2C88DA93EBB6 CRC64;
 Query Match 11.4%; Score 145; DB 1; Length 219;
 Best Local Similarity 24.2%; Pred. No. 0.00012; Mismatches 96; Indels 36; Gaps 8;
 Matches 56; Conservative 43; Mismatches 96; Indels 36; Gaps 8;
 QY 14 SPSKALEPDDKGRVREIAQGLKTSLAAGLGMVPIGIAFGILVIGYEWMAAPFSGIL 73
 DB 3 SLKSGAM-----FKGLVYSPPIYMAIYPAVFTGVALRTLGFSEVAMLASLL 51
 QY 74 IFAGSTEMLVVAVGAAPGAIATLTLVNFHRYA-----PSFLHVVKPIAR 125
 DB 52 IFAGASOFALITL-YQSOLSAIFI-PIPLNLRHIIYSSIIAOKLKLRF-HI----- 101
 QY 126 FYSVFALIDEAVYNTARPGWSAMRLISMQIAFHSVWVGGLTGVAAIALIFPE---IK 182
 DB 102 --SAFGLTDEVFAVSV--NSAENRFLGLGLSISYSAWVGTAAGTGLIDRIVYS 157
 QY 183 GLERFALCSLFTTLTDSCKTKQIPSLLAGLSFTIALVVIPQALPAAL 233
 DB 158 ALVFISIALFVILLPLNKGHRVAAVSGAVLAFLMLTSGVIGIIAAL 208
 RESULT 7
 N01C_PLEBO STANDARD; PRT; 372 AA.
 AC Q00242;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NAD(P)H-quinone oxidoreductase chain 1 (EC 1.6.5.-) (NAD(P)H
 DE dehydrogenase 1, chain 1) (NDH-1, chain 1).
 GN NDHA OR NDH1.
 OS Plectonema boryanum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
 ON NCBI_Taxid=1184;
 RX NCB1
 RP SEQUENCE FROM N.A.

[1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling U.-S.L., Noir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Carnuo A., Uria-Mickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*.";
 RL Nature 397:176-180(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001548; AAD06824.1; -
 CC PIR; F71831; F71831.
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KM Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 130 150 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 SQ SEQUENCE 228 AA; 25844 MW; 5D6EBA86BE2B9744 CRC64;
 Query Match 18.1%; Score 229.5; DB 1; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2.8e-11;
 Matches 61; Conservative 41; Mismatches 95; Indels 9; Gaps 5;
 QY 44 LGMTPIGAFGLVLYQYGEWMAAPLFGSLIFAGSTEMVLYALVGAAPLGAIALTLTLV 103
 Db 20 LGYILMGMTFGWLLAQGGDYKVALFMSLIFYAGIQVVALTLASAQSLNNVIVSLV 79
 QY 104 NERHVFYAFSPFLHVNKPPIARF-VSPFALIDEAVT--AASPAGSAMRLISMQIAF 159
 Db 80 NARQCIALSW-LDFPKTKTKRLPYLALALDDETALNLVAPKGVNETPFMSISILN 138
 QY 160 HSYWVFGLTGVVAIAELIFPEIKGLEFALCSLFTVTLTDSCTKQIPSLLAGISFTIA 219
 Db 139 HSYWVFGLTGVVAIAELIFPEIKGLEFALCSLFTVTLTDSCTKQIPSLLAGISFTIA 219
 QY 220 LVVITPGQALF--ALLIFLGLTLTR 242
 Db 198 CLAPFTEYFLLIALVLMVLAFLTR 223
 RESULT 4
 YH38_HAEIN STANDARD; PRT; 244 AA.
 ID_YH38_HAEIN
 AC P44302;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11738.
 GN H11738.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCB1_TaxID=727;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=95530630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weiman J.F., Phillips R., Spillane R., Hunkeler E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Sprague T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.D., Fritchman J.D., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32846; AAC23382.1; -
 CC PIR; P64041; P64041.
 DR TIGR; H11738; -
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KM Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 SQ SEQUENCE 244 AA; 27360 MW; 7C61B72D4AF16021 CRC64;
 Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4.3e-11;
 Matches 57; Conservative 50; Mismatches 87; Indels 13; Gaps 6;
 QY 45 GMYPIGAFGLVLYQYGEWMAAPLFGSLIFAGSTEMVLYALVGAAPLGAIALTLTLV 103
 Db 30 GFLFLGAYGIYKALGFGFLYPTLMALITYGVSVEFIAGALIAFSPV-SVLTITLMI 88
 QY 104 NERHVFYAFSPFLHVNKPPIARF-VSPFALIDEAVT--VTARPGWSAMRLISMQ 156
 Db 89 SANQIFRGISMLEKIGIHGK--RWYLTITLVDSPELNTWAKIPHLDKGMVFFVS 145
 QY 157 IAFHSYWVFGLTGVVAIAELIFPEIKGLEFALCSLFTVTLTDSCTKQIPSLLAGIS 215
 Db 146 IYHIIYVWGAAAGNLFGLVLPNLKGVFSMTALFLVFAEWMLKQKSHSSILGIGIA 205
 QY 216 FTIALVITPGQALFALLIFLGLTLTR 242
 Db 206 LVFLILIGKXYFLIPTIGIWLITMK 232
 RESULT 5
 YGAZ_ECOLI STANDARD; PRT; 245 AA.
 ID_YGAZ_ECOLI
 AC P76630;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein YGAZ.
 GN YGAZ OR B2682.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCB1_TaxID=562;
 OX [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.M., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of Escherichia coli K-12."
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AE000353; AAC75729.1; -
 CC PIR, C65048; C65048.
 CC DR Ecocyc; E613528; Y9A2.
 CC DR InterPro; IPR004471; Azic.
 CC DR Pfam; PF03591; Azic; 1.
 CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 CC FT TRANSMEM 25 45 POTENTIAL.
 CC FT TRANSMEM 64 84 POTENTIAL.
 CC FT TRANSMEM 110 130 POTENTIAL.
 CC FT TRANSMEM 141 161 POTENTIAL.
 CC FT TRANSMEM 173 193 POTENTIAL.
 CC FT TRANSMEM 206 226 POTENTIAL.
 CC SEQUENCE 245 AA; 26107 MW; 22ACE8AB8D7D651B CRC64;

Query Match 15.5%; Score 197; DB 1; Length 245;
 Best Local Similarity 29.1%; Pred. No. 1.1e-08; Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;

QY 34 OGKLTSLAAGLQYPIGAFGLVIOGYEWMAAPLFGSLIPAGSTEMLVIALVGAAPL 93
 DB 17 EGCKSLPIVIGIYIPAFAGLNAATRLGSPLESVFSCIIYAGASQVITMLAAGSL 76
 QY 94 GALTTLTLVNPFRVYVAPFLAVH---KNPIAFSVFPLIDEAVAVTAAR---PA 145
 DB 77 WIALTLVMDVRAHVLYGSLRSRIIRIQSKTLW--AFGLTDEVRPAATAKVRNR 134
 QY 146 GMSAMRLISMOIAFHSYWF---GGLTGVATAEIIPFRIKLEFALCSLPVTLTLDSCR 201
 DB 135 RMSENMWIGIAFSMSWVFGLTVIGAFSGSGLLGYPAVEALGFMLPALFMSFLASFO 194
 QY 202 TKKQISLLAGISFTIALV-VIPQALPA---ALL--IFGLLT--IRYFELG 247
 DB 195 RKQ-----SLCVPALVGLAGVTLFSPVAILAGVCGCTLALIOAFWQG 240

RESULT 6

YHS5_ARCFU STANDARD; PRT; 219 AA.
 ID YHS5_ARCFU
 AC 028519;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1755.
 GN AF1755.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.,
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus."
 RL Nature 390:364-370(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AE000982; AAB89495.1; -
 CC PIR, B69469; B69469.
 CC DR TIGR; AF1755; -
 CC DR InterPro; IPR004471; Azic.
 CC DR Pfam; PF03591; Azic; 1.
 CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 CC FT TRANSMEM 14 34 POTENTIAL.
 CC FT TRANSMEM 37 57 POTENTIAL.
 CC FT TRANSMEM 123 143 POTENTIAL.
 CC FT TRANSMEM 155 175 POTENTIAL.
 CC FT TRANSMEM 189 209 POTENTIAL.
 CC SEQUENCE 219 AA; 22337 MW; 265A2C88AD93EBB6 CRC64;

Query Match 11.4%; Score 145; DB 1; Length 219;
 Best Local Similarity 24.2%; Pred. No. 0.00012; Matches 56; Conservative 43; Mismatches 96; Indels 36; Gaps 8;

QY 14 SPKALDEPDDKGRYREINQGLKTSIAAGLQYPIGAFGLVIOGYEWMAAPLFGSL 73
 DB 3 SLKSNW-----FRKGLVSPFIYMAIYIPAFTEGVLARTLGSEVENAMASL 51
 QY 74 IFAGSTEMLVIALVGAAPLGAATLTLLVNPFRVYA-----FSPFLHVVKPIAR 125
 DB 52 IFAGASQFALITL-YQSLSIAIFI-PIFNLNHIYSSIIAQKLRFP-HI----- 101
 QY 126 FYSVFALIDEAYAVTAARPAAGSAMRLISMOIAFHSYWFVGLTGVATAEIIPFE--IK 182
 DB 102 --SAFGLTDEVFAVSV--NSAENRFLGLGLSISAMVGTALGVLAGSTLLIDRDVYS 157
 QY 183 GLEFALCSLPVTLTLDSCRTRKQIPSLILAGISFTIALV-VIPQALPAALL 233
 DB 158 ALVFSISALFLVLLPVLKGRHVRAVSGGVALAFHLLNLTISVGIIAAL 208

RESULT 7

NTIC_PLEBO STANDARD; PRT; 372 AA.
 ID NTIC_PLEBO
 AC Q00242;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NAD(P)H-quinone oxidoreductase chain 1 (EC 1.6.5.-) (NAD(P)H
 DE dehydrogenase I, chain 1) (NDH-1, chain 1).
 GN NDH1 OR NDH1.
 OS Plectonema boryanum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
 OX NCBI_TaxID=1184;
 [1]
 RP SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.
 RP MEDLINE=99120557; PubMed=9923682;
 RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deGange B.L., Carmel G.,
 RA Tummino P.J., Carnes A., Urie-Nickelsen M., Mills D.W., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*."
 RL Nature 357:176-180(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azlC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB001548; AAC06824.1; -
 CC PIR: F71831; P71831.
 DR InterPro: IPR004471; AzlC.
 DR Pfam: PF03591; AzlC; 1.
 DR TIGRFAMs: TIGR00346; azlC; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 130 150 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 FT TRANSMEM 228 AA; 25844 MW; SD6EBA86BE29744 CRC64;
 SQ SEQUENCE
 Query Match 18.1%; Score 229.5; DB 1; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2.8e-11;
 Matches 61; Conservative 41; Mismatches 95; Indels 9; Gaps 5;
 QY 44 LGMTPIGAFGLVLYQYGYEMWAPLFGSLIFAGSTEMVLYLVGAAPLGAIALTLTV 103
 DB 20 LGVLTGMTFGMLAACQGYDYKALFMSLFTYAGAIQFVALTLLSAQASIMNVTVSLV 79
 QY 104 NERHVFYAFSPPLHVVKPIARF-YSVFALIDEAVAT---AARPAQMSARLISQIAR 159
 DB 80 NARQTCVLSM-LDRFKTKTKRLPYLAHLTDFALLNLVYAPKKGVNDPMSISLNL 138
 QY 160 HSYWVFGGLTGAVALIPEIKGLEFALCSLFTVTLTDSCKTKQIPLSLLAGLSTIA 219
 DB 139 HSYWVFGSLVSLVSGSHFSFDQGWEEFWMTAIFVLFWMEQYKRNTHKNMVL-GIALAVV 197
 QY 220 LVVIFGQALF---AALLIFLGLTIR 242
 DB 198 CLALFGTGYFLILVLLVNLALILFR 223
 RESULT 4
 YH38_HABIN
 ID YH38_HABIN STANDARD; PRT; 244 AA.
 AC P44302;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11738.
 GN H11738.
 OS Haemophilus influenzae;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxId=727;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=rd / KW20 / ATCC 51907;
 RX MEDLINE=95530630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Pine L.D., Fritchman J.D., Fritchman J.L., Geoghagan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd."
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azlC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32846; AAC2382.1; -
 CC PIR: F64041; F64041.
 DR TIGR: H11738.
 DR InterPro: IPR004471; AzlC.
 DR Pfam: PF03591; AzlC; 1.
 DR TIGRFAMs: TIGR00346; azlC; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 129 159 POTENTIAL.
 FT TRANSMEM 169 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 244 AA; 27360 MW; 7C61B72D4AF16021 CRC64;
 SQ SEQUENCE
 Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4.3e-11;
 Matches 57; Conservative 50; Mismatches 87; Indels 13; Gaps 6;
 QY 45 GMYPIGAFGLVLYQYGYEMWAPLFGSLIFAGSTEMVLT-ALVGAAPLGAIALTLTV 103
 DB 30 GELFGLIAGIYWKALFGFLYPTLMALIVAGSEVFIAAGLILAFSPV-SVLLITLMI 88
 QY 104 NERHVFYAFSPPLHVVKPIARF-YSVFALIDEAVAT---VTAARPAQMSARLISQI 156
 DB 89 SARQIVYGISMLERKGIHIGK--RWYITLVDSFSLNMAKIPRLDGWMTFVS 145
 QY 157 IAFHSYWVFGGLTGAVALIPEIKGLEFALCSLFTVTLTDS-CKTKQIPLSLLAGIS 215
 DB 146 LVYHITWVLAGAMGNLFTVLPFNKGVFSMTALFLVIFANWMLKGSHESSLGLGIA 205
 QY 216 FTVLVVIFGQALFALLIFLGLTIR 242
 DB 206 LVFLITIGKXYFLIPTLIGIWLITWR 232
 RESULT 5
 YGAZ_ECOLI
 ID YGAZ_ECOLI STANDARD; PRT; 245 AA.
 AC P76530;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein YGAZ.
 GN YGAZ OR B2682.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCBI_TaxId=562;
 RN (1)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:16 ; Search time 13.9833 Seconds

(without alignments)
934.659 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269

Sequence: 1 MCKTQIHSLSEVSPSKAL.....LIIFGLLTIRYFPGKAK 251

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	18.9	254	1	AZLC_BACSU
2	235	18.5	228	1	YD31_HELPY
3	229.5	18.1	228	1	YD31_HELPY
4	227.5	17.9	244	1	YH38_HABIN
5	197	15.5	245	1	YH55_ARCFU
6	145	11.4	219	1	YH55_ARCFU
7	110.5	8.7	372	1	NUIC_PLEBO
8	101.5	8.0	498	1	NU4C_PLEBO
9	101.5	8.0	746	1	CLC5_HUMAN
10	100.5	7.9	746	1	CLC5_MOUSE
11	100.5	7.8	746	1	CLC5_MOUSE
12	98.5	7.8	643	1	SL12_HUMAN
13	98	7.6	760	1	CLC4_HUMAN
14	96.5	7.7	416	1	RHCE_HUMAN
15	95	7.5	747	1	CLC4_MOUSE
16	94	7.4	567	1	PROL_TLREN
17	93	7.3	266	1	PTNC_ECOLI
18	92	7.2	447	1	CLC4_RAT
19	91.5	7.2	764	1	GALP_ECOLI
20	91.5	7.2	695	1	YD38_YEAST
21	91	7.2	624	1	COX1_BACPF
22	90.5	7.1	226	1	ATP6_RHTUN
23	90.5	7.1	267	1	UPK2_MOUSE
24	90.5	7.1	351	1	YB76_MOUSE
25	90.5	7.1	393	1	SETB_ECOLI
26	90	7.1	220	1	YICG_HABIN
27	90	7.1	416	1	RHD_HUMAN
28	90	7.1	416	1	RHD_HUMAN
29	90	7.1	528	1	RHLD_GORGO
30	90	7.1	528	1	TH12_TRYBA
31	90	7.1	528	1	TH2A_TRYBA
32	89.5	7.1	319	1	EMA3_CABEL
33	89.5	7.1	381	1	PSTC_ECOLI

ALIGNMENTS

RESULT 1	AZLC_BACSU	STANDARD;	PRT;	254 AA.
AC	007942:			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Branched-chain amino acid transport protein azlc.			
GN	AZLC OR BSU26710.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / BGSC1A1;			
RX	MEDLINE=97431495; PubMed=9287000;			
RA	Beltsky B.R., Gustafsson M.C.U., Sonenshein A.L., von Wachenfeldt C.;			
RT	"An lrp-like gene of Bacillus subtilis involved in branched-chain amino acid transport."			
RL	J. Bacteriol. 175:5448-5457 (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=97453479; PubMed=9308178;			
RA	Kunert F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RT	Azevedo V., Bortier M.G., Bessieres P., Bolotin A., Borchert S.,			
RL	Boriss R., Bortier L., Brane A., Braun M., Bisgelli S.C., Bron S.,			
RA	Broillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,			
RT	Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,			
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,			
RT	Enutan K.D., Errington J., Fabre C., Ferrari E., Foulger D.,			
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,			
RT	Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,			
RA	Giuseppe G., Guy B.J., Haga K., Haleth J., Harwood C.R., Henaut A.,			
RT	Halbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,			
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,			
RT	Kobayashi Y., Koester P., Koningsen G., Krogh S., Kumano M.,			
RA	Kurita K., Lapidus A., Lardinois S., Lauder J., Lazarevic V.,			
RT	Lee S.W., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,			
RA	Medina N., Meilieu R.P., Mizuno M., Mosati D., Nakai S., Nodack M.,			
RT	Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,			
RA	Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,			
RT	Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,			
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,			
RT	Sato T., Scanlan B., Schleich S., Schroeter R., Scoffone F.,			
RA	Sekiguchi J., Sekowska A., Serror S.J., Serror F., Shin B.S., Soldo B.,			

34	89	7.0	428	1	BCH2_RHOCA
35	88.5	7.0	226	1	ATP6_PHOV
36	88.5	7.0	517	1	COX1_PISOC
37	88.5	7.0	760	1	CLC3_MOUSE
38	88.5	7.0	760	1	CLC3_MOUSE
39	88.5	7.0	760	1	CLC3_MOUSE
40	88.5	7.0	760	1	CLC3_MOUSE
41	88.5	7.0	762	1	CLC3_MOUSE
42	88	6.9	144	1	BDHC_PSERE
43	88	6.9	297	1	BACH_HALS4
44	88	6.9	362	1	NUIC_MAIZE
45	88	6.9	499	1	NU4C_MARPO

P26171	rhodobacter
O00521	phocavitul
P25001	pisaster oc
O91279	cavia porce
P51791	mus musculu
O18894	oryctolagus
P51792	rattus norv
P51790	homo sapien
O8ghn3	pseudomonas
O93741	halobacteri
P25706	zea mays (m
P06263	marichantia

RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassart A.,
RA Viart A., Wandt R., Wedler E., Wedler H., Wetzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RA "the complete genome sequence of the Gram-positive bacterium *Bacillus*
RT subillus".
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: Belongs to the azic family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y1043; CAAT1940.1; -
DR EMBL: U93876; AAB80901.1; -
DR EMBL: Z59117; CAB14612.1; -
DR PIR: G69592; G69592.
DR Subtilist; BG11915; azic.
DR InterPro: IPR004471; Azic.
DR Pfam: PF03591; Azic; 1.
DR TIGRFAMs: TIGR00346; azic; 1.
DR Transport; Amino-acid transport; Transmembrane; Complete proteome.
KM TRANSMEM 33 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
FT TRANSMEM 220 240 POTENTIAL.
SO SEQUENCE 254 AA; 28365 MW; 1463E390B0054B45 CRC64;

Query Match 18.9%; Score 240; DB 1; Length 254;
Best Local Similarity 30.5%; Pred. No. 4.7e-12;
Matches 62; Conservative 44; Mismatches 89; Indels 8; Gaps 6;

QY 45 GWYPIGAFGLVITYGYEWMAAPLPSGLIFAGSTEMVIALVGA-APLGAIALTLTV 103
DB 42 GFLFLIAIGIFMHSIGFSAIPIIMISFIFGSMFVANPLLGAPNMAFL-ITLV 100
QY 104 NFRHVFAPSPPLHVVKNPIARFYSVFALIDEAVN--TAAPAGW-SMRLISMOIAF 160
DB 101 NARHLFYGISMLDKYRGTKKKLYLIFGMCDSFSINYANVPANVDKGMFETLLNH 160
QY 161 SYWVFGGLGVATAIELIPEIKGLSEFALCSLFTVTLTLDSCRTKQIPSLILAGLSFTIA 220
DB 161 LVVAVAGALIGTIGSVKNTGELDVTMTALFIVIFIEGMMEKHYS-ALVGLGLSVAS 219
QY 221 VVIPG--QALFALLIFGLTIT 241
DB 220 LILFGGNQFIIPAMLAITGLTV 242

RESULT 2
ID YD31_HELPY STANDARD; PRT; 228 AA.
AC 025889;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein HP1331.
GN HP1331.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_Taxid=210;
RP SEQUENCE FROM N.A.

RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=973994467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khaliq H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.,
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT pylori".
RL Nature 388:539-547(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: Belongs to the azic family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000634; AAD08372.1; -
DR PIR: C64686; C64686.
DR TIGR: HP1331; -
DR InterPro: IPR004471; Azic.
DR Pfam: PF03591; Azic; 1.
DR TIGRFAMs: TIGR00346; azic; 1.
KM Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 130 150 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
SO SEQUENCE 228 AA; 25854 MW; D8EFA6B68376130 CRC64;

Query Match 18.5%; Score 235; DB 1; Length 228;
Best Local Similarity 33.0%; Pred. No. 1.1e-11;
Matches 70; Conservative 36; Mismatches 92; Indels 14; Gaps 7;

QY 44 LKWYPIGAFGLVITYGYEWMAAPLPSGLIFAGSTEMVIALVGAAPLGAIALTLTV 103
DB 20 LGVILMGMTEGMLLVGGGYDYKVALFMSLFTYGAQVFAITLISAQASIMNVIVSLIV 79
QY 104 NFRHVFAPSPPLHVVKNPIARF-YSVFALIDEAVN--TAAPAGW-SMRLISMOIAF 159
DB 80 NARQCYALSM-LDRFNNTWRLPYLAHALTDTFALNLVYAREGVSEKQIPSLISLN 138
QY 160 HSWVFGGLGVATAIELIPEIKGLSEFALCSLFTVTLTLDSCRTKQIPSLILAGLSFTIA 219
DB 139 HSWIFGSLVGSVGHFSFDTQGMETVMTALFIVIFEMEG--YKSTTNKKNMGLVIA 195
QY 220 LVVIPGALPAA--LIFGLLTIIFYFLGK 248
DB 196 VVCL---ALPTEYFELIALVNLVLMLEPRK 224

RESULT 3
ID YD31_HELPY STANDARD; PRT; 228 AA.
AC 092403;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein UHP1251.
GN UHP1251.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_Taxid=85963;
RP SEQUENCE FROM N.A.

Tue Apr 20 06:47:18 2004

us-10-608-504-3.rpr

Page 6

Qy 28 RRYIAQCKTSLAAGLGMYPICIAFGILVIOGYEMWAPLFGSLIFAGSTEMLVIALV 87
Db 11 RETDMLRGAAIGSSIAIGYIPASLTFFGLAKGTGLTLETIAMSLEFYAGASQYALTLT 70
Qy 88 VGAAPLGAIALTTLLVNFPHVFAFSFPLHV-KNPIARFYSVFALIDEAVAVTAARPAG 146
Db 71 AIGTGIEIILTTFTVINIHLMSASIRAKMEDTHPVKALITAFGITDEVFALVTSQDR 130
Qy 147 WSAWRLISMOIAFHSYWFGLTGVAIAELIPPEI-KGLEFALCSLFTTLTIDSCRTKKQ 205
Db 131 LTNGFVIGVAVIAYVSWVHSVAVGYVGSALPATLOQGMGVALYAMFIALLIPSVKHS 190
Qy 206 IPBLLAGLSFTIALVVIPOALFAL-----LIFGLLTRYFPLGKAK 251
Db 191 V--LILAG--TVALL-----NGLFSLFPSPGWSIILATLILASVGYRIGKAK 233

Search completed: April 16, 2004, 06:54:24
Job time : 21.5766 secs

C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 03-Jun-2002
 C:Accession: H83390
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Braham, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Ladbis, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: H83390
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-252 <STO>
 A:Cross-references: GB:AE004630; GB:AE004091; NID:g9948041; PIDN:AG05427.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2039
 C:Superfamily: hypothetical protein b2682

Query Match 13.8%; Score 175; DB 2; Length 252;
 Best Local Similarity 24.2%; Pred. No. 1.3e-07;
 Matches 52; Conservative 47; Mismatches 104; Indels 12; Gaps 6;

29 RREINQGLKSLAAGLGMVPIGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMVIALV 88
 24 RQAFKRGAVAILPLSLAVAPWGLAGSMATEANTLPABQGLSAIVFQAQVAIGMLK 83
 89 GAAPGALATTLVNFHVFYAFSPPLHVNKNPIARFVS---PALIDEAVAVTAAP- 144
 84 GGAGFSSIFFTLLTSLTSLTGLTGMG--LRPVLSPLRGMRIGLGLTDEFALASQDR 141
 145 AGKSAARLISMOIAFHSTYVFGSLGVAIAELIP-PEIKGLEFALCSLFVTLTDSCTK 203
 142 RNFNRVYALGVGLTFYIANLFTLGLGSLIGLEHGLDPSIAATFIALVAPLV-- 199
 204 KOIPSHLAGLSFTIALVVTFCQALFAALLIFGL 238
 200 -NVPTIVCAVATSLFCGVLFSHNQ--WSAVALVLAGL 231

RESULT 13
 AF0396
 probable amino acid transporter YPO3264 [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 01-Mar-2002
 C:Accession: AF0396
 R:Fairhall, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11563360
 A:Accession: AF0396
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <RUR>
 A:Cross-references: GB:AL590842; PIDN:CA092498.1; PID:g15981198; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO3264
 C:Superfamily: hypothetical protein b2682

Query Match 13.8%; Score 175; DB 2; Length 257;
 Best Local Similarity 25.8%; Pred. No. 1.3e-07;
 Matches 56; Conservative 38; Mismatches 93; Indels 30; Gaps 6;

34 OGKTSIAAGLGMVPIGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMVIALVGAAPL 93
 20 EGTIDSLPIVIGLTPAFAGLSAVLGGTTPLESIFSCITVAGAGQFVITALLSGMGL 79
 94 GAIALTTLLVNFHVFYAFSPPLHVNKNPIARFVS-----VFALIDEAVAVTAAR--- 143
 80 WVASLITVMAVDVHIILYG-----PALKRIILAKLSKKTALVAFGLTDEVFAAATTKMK 134

QY 144 -PAGSARLISMOIAFHSTYVFGSLGVAIA-----ELIPFIEKLEFALCSLFVTLTLD 198
 DB 135 DQRRSENMWGLIAVTSWTLGTAIGAMPNGPLENPALIASLSFVLPALFSLFLA 194
 QY 199 SCRTKQI08BLIAGS-----FTIALVVIPOGA 227
 DB 195 S-FKQYSITVYASISGALLGVLLFSTIVAILAGIA 229

RESULT 14
 E86826
 amino acid permease ynfD [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
 C:Species: *Lactococcus lactis* subsp. *lactis*
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: E86826
 R:Bolet, A.; Winkler, P.; Mager, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Enrli Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
 A:Reference number: A86825; MUID:21235186; PMID:11337471
 A:Accession: E86826
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-235 <STO>
 A:Cross-references: GB:AE005176; PID:g12724620; PIDN:AAK05711.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ynfD

Query Match 13.8%; Score 174.5; DB 2; Length 235;
 Best Local Similarity 25.5%; Pred. No. 1.3e-07;
 Matches 61; Conservative 47; Mismatches 92; Indels 39; Gaps 9;

34 OGKTSIAAGLGMVPIGIAFGLLVIOGYEWMAAPLFGSLIFAGSTEMVIALVGAAPL 93
 9 OGKIDMPVTFVIGYIGIAFGMTGHSFVWVILLISLIVAGSAQFTVWSMLATHSPF 68
 94 GAIALTTLLVNFHVFYAFSPPLHVNKNPIARFVSFALI-DEAVAV----- 139
 69 MSIVLSFVLYNSRIILMSMTATSYFNKESILKNIILGTLTDESFLGMKXNYTGKLN 128
 140 ---TARPAGKSAARLISMOIAFHSTYVFGSLGVAIAELIPPEIKGLEFALCSLFV-- 193
 129 FSWFNANLALAYVMAAS-----AIGALLGMLAN--P-EKIGLGFVIAIAFTGL 176
 194 -TTLTDSCTRKQI08BLIAGS-----FTIALVVIPOG--ALFALD-IFGLITIRYFF 245
 177 LYQLDISDKTLGLMLQLVMTGIVLVLFYFGLIFLPSNLVLFTVLLACALGVGVKVEFF 235

RESULT 15
 F84013
 branched-chain amino acid transporter BH2910 [imported] - *Bacillus halodurans* (strain C-
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Mar-2002
 C:Accession: F84013
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: F84013
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-237 <STO>
 A:Cross-references: GB:AB001517; GB:BA000004; NID:g10175500; PIDN:BA06629.1; GSPDB:GN00
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH2910
 C:Superfamily: hypothetical protein b2682

Query Match 13.8%; Score 174.5; DB 2; Length 237;
 Best Local Similarity 23.7%; Pred. No. 1.3e-07;
 Matches 55; Conservative 48; Mismatches 112; Indels 17; Gaps 6;

QY 88 VCAALPGALATLTLLVNRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAY 137
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 66 THPASMWTLGFPALLVNRHVLMSAS-----IAGKLDDFKGMRKXKAMLVLTDESW 116

QY 138 AVTAAR-----PAGMSANRLISMOIAFHSYVFGSLTGVAIAELIPEIKGLEFALCSL 191
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 117 ALABRVIAGRLTAFAFGAALSIYLVNLTATLGAALGAVVGDV---SVIGLDPAFPAY 173

QY 132 FVTLLDSCRTKQIPSLILAGLSFTIALVVIPOALPAA 231
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 174 FIVLMGFVKRGRETVGLVILASASACUTLALVPGAWYIAA 213

RESULT 9

B97492
 azic family protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: B97492
 A:Accession: B97492
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Matkelsz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:21608551; PMID:11743194
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Query Match 14.3%; Score 181; DB 2; Length 234;
 Best Local Similarity 25.0%; Pred. No. 3.6e-08;
 Matches 55; Conservative 46; Mismatches 91; Indels 28; Gaps 4;

QY 28 RRYEINAGLKTSLAGIGWPIGIAFGLVIOGYEMWAAPIPSGLIPAGSTEMVIALY 87
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 6 KRAELIAGLSAAPPLVAVPVGVAIGKLSPLPSLSLVFAGSGQFVAMDLM 65

QY 88 VCAALPGALATLTLLVNRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAY 137
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 66 THPASMWTLGFPALLVNRHVLMSAS-----IAGKLDDFKGMRKXKAMLVLTDESW 116

QY 138 AVTAAR-----PAGMSANRLISMOIAFHSYVFGSLTGVAIAELIPEIKGLEFALCSL 191
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 117 ALABRVIAGRLTAFAFGAALSIYLVNLTATLGAALGAVVGDV---SVIGLDPAFPAY 173

QY 132 FVTLLDSCRTKQIPSLILAGLSFTIALVVIPOALPAA 231
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 174 FIVLMGFVKRGRETVGLVILASASACUTLALVPGAWYIAA 213

RESULT 10

G82388
 azic family protein VCA1002 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 01-Mar-2002
 C:Accession: G82388
 R:Heideberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.; Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; PMID:20406833; PMID:10952301
 A:Accession: G82388
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-239 <HEI>
 A:Cross-references: GB:AE004427; GB:AE003853; NID:G9658442; PIDN:AAF96898.1; GSPDB:GN001

Query Match 14.3%; Score 181; DB 2; Length 234;
 Best Local Similarity 25.0%; Pred. No. 3.6e-08;
 Matches 55; Conservative 46; Mismatches 91; Indels 28; Gaps 4;

QY 28 RRYEINAGLKTSLAGIGWPIGIAFGLVIOGYEMWAAPIPSGLIPAGSTEMVIALY 87
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 6 KRAELIAGLSAAPPLVAVPVGVAIGKLSPLPSLSLVFAGSGQFVAMDLM 65

QY 88 VCAALPGALATLTLLVNRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAY 137
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:gl156113; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Db 66 THPASMWTLGFPALLVNRHVLMSAS-----IAGKLDDFKGMRKXKAMLVLTDESW 116

A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA1002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Query Match 14.3%; Score 181; DB 2; Length 239;
 Best Local Similarity 23.6%; Pred. No. 3.7e-08;
 Matches 48; Conservative 47; Mismatches 104; Indels 4; Gaps 4;

QY 34 QGKTSIAGLGVPIGIAFGLVIOGYEMWAAPIPSGLIPAGSTEMVIALVGAAL 93
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 20 QGTAMLPISIALVPMGLAGSPFAIEAGLSVIESQALSAVLAQAQVAIGMFTGAGL 79

QY 94 GATLTLLVNRHVFYAFSPPLHVKNPIA--RFYSVFALIDEAVTAAR--PAGMSANR 151
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 80 LSLIATFPITSHRFLYSVMSKSIPLRMLTIGFILLDELPAICGASDKQFNRY 139

QY 152 LISMOIAFHSYVFGSLTGVAIAELIPEIKGLEFALCSLFTVTLTDSCTKQIPSL 210
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 140 ALGAGLSFYLIWMLASLVAGSYLPDLNOMGLEPAAVATFAIVIPNIKSWPVLISVL 199

QY 211 LA-GLSTIALVVIPOALPAA 232
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 200 TALVLSVLTVMGIESLMPASI 222

RESULT 11

G95016
 conserved hypothetical protein SP0146 [imported] - Streptococcus pneumoniae (strain TIGR)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: G95016
 R:Heideberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.; Nature 406, 477-483, 2000
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; PMID:21357209; PMID:11463916
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Query Match 14.0%; Score 177.5; DB 2; Length 218;
 Best Local Similarity 26.6%; Pred. No. 6.7e-08;
 Matches 63; Conservative 36; Mismatches 82; Indels 39; Gaps 11;

QY 40 LAAGLWYPIGIAFGLVIOGYEMWAAPIPSGL-----IFGSTEMVIALVGAALPGA 95
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 1 MPTALGVYSIGLAGIIGAPY-----VTPVEMGMSLFVYAGSAQFALAIIVQAPVAA 55

QY 96 IALTLLVNRHVFYAFSPPLHVKNPIA--RFYSVFALIDEAVTAAR--PAGMSANR 146
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 56 IATMFTLINLR---LFLSLHASTYFRHTLWNIGMSILTBETTVLNGELAHDKV 111

QY 147 ---WSANRLISMOIAFHSYVFGSLTGVAIAELIPEIKGLEFALCSLFTVTLTDSCT 201
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 112 NEMWNGNNLNSYVA---WFGVGVGALGLLPNPEIFGLDPALVGMFTGIFASOPQI 167

QY 202 TKQIP-----SLIAGLSFTIALVVIPOALPAA 233
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:gl4971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 A:Gene: SP0146

Db 168 MORRIPVNNLIIILAVAVSPFLITVMSQSLAVIFATLL 207

RESULT 12

H83390
 hypothetical protein PA2039 [imported] - Pseudomonas aeruginosa (strain PA01)

Db 101 NARHLYGISMIDKRYGTGKKLYLIFGWCDESFISINTANVPANVDKGMFVTLTNN 160
 QY 161 SYWVVGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTKQKQIPSLLAGLSFTIA 220
 Db 161 LTVWVGAAIGIFGSGYVKNTEGDDVMTALFIVLFIEQMKKEKHYS-ALTGLGLSYAS 219
 QY 221 VVIPG--QALFPAALLIFGLLTI 241
 Db 220 LILFGNGQFIIPAMLAITGLVTL 242

RESULT 2

conserved hypothetical integral membrane protein HP1331 - Helicobacter pylori (strain 26 C64686)
 C/Species: Helicobacter pylori
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jun-2000
 C/Accession: C64686
 R/Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C. A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A/Reference number: A64520; MUID:97394467; PMID:9252185
 A/Accession: C64686
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-228 <TOM>
 A/Cross-references: GB:AE000634; GB:AE000511; NID:G2314489; PIDN:AAD08372.1; PID:G231449
 C/Superfamily: hypothetical protein b2682

Query Match 18.5%; Score 235; DB 1; Length 228;
 Best Local Similarity 33.0%; Pred. No. 9, 2e-13;
 Matches 70; Conservative 36; Mismatches 92; Indels 14; Gaps 7;

QY 44 LGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 20 LGYILMGMTGMLLVQGGYQYKVALPMSLFYAGAVQFAVATLLSAQASLMNVIVSLIV 79
 QY 104 NFRHVFYAFSPFLHVNKNPIARF-YGVFALIDEAVV--TAARPAQMSAMRLI-SMQIAF 159
 Db 80 NARQTCYALSM-LDRFNKTKWRLPYLAHALTDFTFALLNLVAPKGVSKDFIFISISLN 138
 QY 160 HSFWVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTKQKQIPSLLAGLSFTIA 219
 Db 139 HSWVIFSLVGSIVGSHFSFDQGMFVMTALFIVLFMEQ--YKRTNHKNAMIGIVIA 195
 QY 220 LVVIPGQALFPA--LILFGLTIRYFPLGK 248
 Db 196 VVCL--ALFTEYFLLIALVLMVLAFLMRK 224

RESULT 3

F71831
 hypothetical protein jhp1251 - Helicobacter pylori (strain J99)
 C/Species: Helicobacter pylori
 A/Variety: strain J99
 C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
 C/Accession: F71831
 R/Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, B.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
 A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path A/Reference number: A71800; MUID:99120557; PMID:9923682
 A/Accession: F71831
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-228 <ARN>
 A/Cross-references: GB:AE001548; GB:AE00439; NID:G4155845; PIDN:AAD06624.1; PID:G415585
 A/Experimental source: strain J99
 C/Genetics:
 A/Gene: jhp1251
 C/Superfamily: hypothetical protein b2682

Query Match 18.1%; Score 229.5; DB 2; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2, 7e-12;
 Matches 61; Conservative 41; Mismatches 95; Indels 9; Gaps 5;

QY 44 LGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 20 LGYILMGMTGMLLVQGGYQYKVALPMSLFYAGAVQFAVATLLSAQASLMNVIVSLIV 79
 QY 104 NFRHVFYAFSPFLHVNKNPIARF-YGVFALIDEAVV--TAARPAQMSAMRLISMQIAF 159
 Db 80 NARQTCYALSM-LDRFNKTKWRLPYLAHALTDFTFALLNLVAPKGVNETDEMFISISLN 138
 QY 160 HSFWVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTKQKQIPSLLAGLSFTIA 219
 Db 139 HSWVIFSLVGSIVGSHFSFDQGMFVMTALFIVLFMEQYKSNVNHKNAML-GIAIAYV 197
 QY 220 LVVIPGQALF--AALIFGLTIR 242
 Db 198 CLALFTEYFLLIALVLMVLAFLMR 223

RESULT 4

F64041
 hypothetical protein H11738 - Haemophilus influenzae (strain Rd KM20)
 C/Species: Haemophilus influenzae
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
 C/Accession: F64041
 R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J. R/Fleischmann, R.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C. D.M.; Brannon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrman, J.L.; Geoghegan, N.S.W. Science 269, 496-512, 1995
 A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A/Reference number: A64000; MUID:95350630; PMID:7542800
 A/Accession: F64041
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-244 <TIGR>
 A/Cross-references: GB:U32846; GB:L42023; NID:G3212237; PIDN:AC23382.1; PID:G1574596; 1
 C/Superfamily: hypothetical protein b2682

Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4, 3e-12;
 Matches 57; Conservative 50; Mismatches 87; Indels 13; Gaps 6;

QY 45 GMYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 30 GFLPLGIAVGYMALGFGFLYPTLMALLTAGSVERIAGALAPSPSI-SVLLITIMI 88
 QY 104 NFRHVFYAFSPFLHVNKNPIARFYSVALIDEAVV--VTAAAPAGMSAMRLISMQ 156
 Db 89 SARQIFYGISMLEKYGHIHCK--RWLYITTLVDEBSFSLNYAKIPDHLDRGYMFEVS 145
 QY 157 IAFHSYVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDS-CTKQKQIPSLLAGLS 215
 Db 146 LYLHLYVTLGAAMNLRCTVVPNMLKGVFMTLFLVIFPKNMLKXSHSSSLGIGIA 205
 QY 216 FTIALVIVPGQALFPAALLIFGLLTI 242
 Db 206 LVFLLITIGKEYFLIPTLIGIWLITMR 232

RESULT 5

C85916
 hypothetical protein 23983 [imported] - Escherichia coli (strain O157:H7, substrain EDL)
 C/Species: Escherichia coli
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
 C/Accession: C85916
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grobeck, E.J.; Davis, N.W.; Linn, A.; Dimalanta, E.; Potamousis, K.; Agodaca, Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 19.5766 Seconds

(without alignments)
1233.313 Million cell updates/sec

Title: US-10-608-504-3
Perfect score: 1269
Sequence: 1 MORTQEIHSLEVSFKAL.....LIFLGILTRFYFLGKAAK 251

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR.78.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	240	18.9	254 1	branched-chain am
2	235	18.5	1	conserved hypothet
3	229.5	18.1	228 1	hypothetical prote
4	227.5	17.9	244 1	hypothetical prote
5	198	15.6	245 2	hypothetical prote
6	198	15.6	245 2	hypothetical prote
7	197	15.5	245 1	hypothetical prote
8	181	14.3	234 2	branched-chain am
9	181	14.3	234 2	branched-chain am
10	181	14.3	239 2	azlc family protei
11	177.5	14.0	218 2	conserved hypothet
12	175	13.8	252 2	hypothetical prote
13	175	13.8	252 2	hypothetical prote
14	174.5	13.8	235 2	probable amino aci
15	174.5	13.8	235 2	amino acid permeas
16	169	13.3	235 2	branched-chain am
17	166	13.1	235 2	transport protein
18	163.5	12.9	219 2	probable exported
19	147	11.6	231 2	hypothetical prote
20	145.5	11.5	240 2	conserved hypothet
21	145.5	11.5	240 2	azlc family protei
22	145	11.4	219 1	conserved hypothet
23	140.5	11.1	246 2	branched-chain am
24	140	11.0	160 2	hypothetical prote
25	137	10.8	235 2	azlc family protei
26	115	9.1	140 2	azlc-related prote
27	113.5	8.9	477 2	probable amino aci
28	112	8.8	379 2	Na+/H+ antiporter
29	112	8.6	490 2	probable sugar tra

30	110.5	8.7	372 2	J02135	NADH2 dehydrogenas
31	110.5	8.7	508 2	AH2688	Na+/H+ antiporter
32	110.5	8.7	525 2	C97470	Probable NADH dehy
33	108.5	8.6	334 2	B96997	permease (imported
34	108	8.5	256 2	T35133	hypothetical prote
35	107	8.4	399 2	AH2542	hypothetical prote
36	106.5	8.4	487 2	E96782	hypothetical prote
37	103	8.1	408 2	B83333	probable MFS trans
38	103	8.1	408 2	B87889	NADH dehydrogenase
39	102	8.0	481 2	C95920	hypothetical membr
40	101.5	8.0	417 2	T51467	glucose 6 phosphat
41	100	7.9	260 2	E83174	hypothetical prote
42	100	7.9	541 2	AH2909	hypothetical prote
43	100	7.9	541 2	F97684	thiamin transport
44	99.5	7.8	455 2	A75121	hypothetical prote
45	99	7.8	233 2	F97382	probable transport

ALIGNMENTS

RESULT 1

branched-chain amino acid transport azlc - Bacillus subtilis

C/Species: Bacillus subtilis
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C/Accession: G69592; T44777

R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, K.D.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
Jech, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A/Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Seror
Akenuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Zundstein, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MID:198044033; PMID:9384377

A/Accession: G69592
A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
A/Residues: 1-254 <BFL>
A/Cross-references: GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CA814612.1; PID:G2635116
A/Experimental source: strain 168
R/Belitsky, B.R.; Gustafson, M.C.U.; Sonenshein, A.L.; von Wachenfeldt, C.
J. Bacteriol. 179, 5448-5457, 1997

A/Title: An lrp-like gene of Bacillus subtilis involved in branched-chain amino acid tra
A/Reference number: Z22837; MID:97431495; PMID:9287000

A/Accession: T44777
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-254 <BFL>
A/Cross-references: EMBL:Y11043; NID:G1926275; PIDN:CA81940.1; PID:G1926281
A/Experimental source: strain 1A1
C/Genetics:
A/Genes: azlc
C/Suprafamily: hypothetical protein b2662

Query Match 18.9% Score 240; DB 1; Length 254;
Best Local Similarity 30.5% Pred. No. 3.9e-13;
Matches 62; Conservative 44; Mismatches 89; Indels 8; Gaps 6;

QY 45 GMYPIGIAFLVYQGYEWMAAPLPSGLIPAGSTEMLVTAIVGA-APLGAIAITLLV 103
DB 42 GFLPLGIAYGIFMHSIGFSAIYPIIMSPMIFAGSMFEVAFNPLGAFNMNALFL-TLV 100
QY 104 NFRVVFAPSPFLHVNKNPIARFYSFALIDEVAA--TAAEPAGN-SAMRLISQIAFH 160